1964

A Study of the Methods of Interpreting Measures of Achievement in Relation To Measures of Capacity

John Arthur Russell
Loyola University Chicago

A STUDY OF THE METHODS OF INTERPRETING MEASURES OF ACHIEVEMENT IN RELATION TO MEASURES OF CAPACITY

by

John A. Russell

A Dissertation Submitted to the Faculty of the Graduate School of Loyola University in Partial Fulfillment of the Requirements for the Degree of Doctor of Education

June

1964

LIFE

John A. Russell was born in Chicago, Illinois, on May 14, 1930.

He was graduated from Saint Ignatius High School, Chicago, 1948. His undergraduate work was done at the University of Illinois and Loyola University where he received his Bachelor of Science in Social Science, June, 1952. In 1954 he earned a Master of Arts degree from the latter university.

Since 1952 Mr. Russell has been employed by the Chicago Board of Education. From 1952 to 1954 he served as a teacher at Fenger High School. In 1954 he was assigned to Hyde Park High School where he served as a teacher, grade advisor and placement counselor. In September, 1958, he was assigned to the Bureau of Pupil Personnel Services where he worked as a supervisor and assistant director. He was assigned to Horace Mann Elementary School as a principal in September, 1962. In addition to the above experiences he has been employed by Chicago Teachers College and Saint Xavier College as a part time instructor in undergraduate as well as graduate courses in education.

During the summer of 1958 he was responsible for preparing a curriculum course of study in the area of Civics. He

was also responsible for a study guide to be used by high school counselors for a course in Group Guidance. In the summer of 1960 he was responsible for the preparation of a pamphlet "Plotting Your Program" which is used by freshmen in the Chicago Public High Schools.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

During the past eight years, the early identification and education of the "gifted" and talented student has received unprecedented interest and emphasis in American education. Although prior to this time there have been periods of interest, this recent emphasis has taken on a "mystical cloak" of nationalism. It is not only the waste of an individual's talents and abilities, but a loss to society and in particular to the nation which is being recognized. The effects and in some ways the cause of this national concern can be seen in the number of special programs and reports which have been undertaken. Prominent among any list of such programs and reports would be: the National Defense Education Act; the National Merit Scholarship Corporation; Manpower and Education, a publication of the Educational Policies Commission; the North Central Association of Colleges and Secondary Schools, Project on Guidance and Motivation of Superior and Talented Students; the Rockefeller Brothers Fund Report on the Pursuit of Excellence; the books by James B. Conant; and the Talent Preservation High Horizons Projects of New York City. Others could be listed, but those mentioned give some indication of the present nationwide attention which is

being given to the identification and education of the gifted.

Each of the preceding reports and projects has pointed out the need and importance of early identification, in order that proper motivation and education may be insured. One of the most frequently utilized methods of identification has been intelligence tests and achievement tests. It seems almost ironical that this emphasis which has been placed upon the identification of the gifted, should also be the same period that the greatest strides in test scoring have been made. It has been approximated that in 1960, one hundred and thirty million tests were administered and scored.[1] Technological advancements in test scoring have reduced, if not eliminated, the objections that a testing program is a burden upon the teachers' and other school staff's time. The scoring of 500,000 tests, National Merit Participation for 1961, would have been out of the question in previous years. It is now possible to carry out large scale testing programs which would have been rejected because of the time consuming task of scoring and reporting.

Admittedly, schools have used tests for a variety of purposes other than the identification of the gifted. Classifi-

---

1 Ruth Barry and Beverly Wolf, An Epitaph For Vocational Guidance, New York, 1962, p. 27.

cations, grouping, diagnosis, growth and individual guidance are but a few of the ways in which tests have been utilized in a school situation. However, it appears almost inevitable that when intelligence tests and achievement tests are administered to a group or an individual pupil, that some method be utilized to relate these two types of tests, in an effort to evaluate achievement in terms of capacity. Thus the practice of relating ability and achievement have taken on increased importance and frequency as a result of the greater number of tests which are being given, as well as the interest in seeing that the able student is achieving in keeping with his abilities.

These factors: (1) interest in and emphasis on the identification of the gifted, (2) increased use of tests, (3) efforts to relate intelligence and achievement, have led this writer to propose the following investigation.

The Purpose of the Study

It is the purpose of this research to examine three aspects of the problem:

1. What are some of the methods which have been utilized to relate measures of intelligence to measures of achievement?
2. What are some of the current practices of relating intelligence and achievement tests in

public school systems with enrollments over 200,000?

3. What are the limitations of the practice of relating intelligence and achievement test results through the use of an age-grade expectancy table as is done in Chicago?

Procedure

The first phase of this problem, a critical examination of the literature relevant to the topic, will be treated in the traditional manner. Particular care, however, will be taken to present this literature in its historical sequence so that it might be seen in the perspective from which it evolved. Although current literature about a topic is generally a reflection of current practice, experimentation and thinking, it was felt that perhaps some practices are being used which have been shown by the research to have limitations. That is, a particular technique or method of relating test results is possibly still being utilized in spite of the shortcomings which have been pointed out by the research.

It is for this reason that the second aspect, a survey of current practices as revealed by a questionnaire sent to fifty-six cities with populations over 200,000 was undertaken. (See Appendix I) An examination of this questionnaire, discussed

more fully in Chapter III, reveals an intended effort to make it brief and unstructured. This was done so that the person responding would not be influenced by what he might consider to be good practice.

The final phase of the problem examined in this research, is a critical analysis of a practice utilized by the Chicago Public Schools, to relate measures of intelligence and achievement. This phase of the research analyzes the intelligence and achievement test data of twenty thousand 6A and twelve thousand 8B Chicago Public School pupils who took the California Short Form Test of Mental Maturity, and the Metropolitan Achievement Battery in the Spring of 1961. For some time the Chicago Public Schools have compared achievement in terms of mental ability through the use of the mental age grade expectancy (M.A.G.E.).[2] This concept would have us reduce mental age and grade placement test scores to an equivalent basis through the use of an expectancy table.[3] The question arises as to how realistic and accurate such a practice is. The writer, in an effort to examine this point, had prepared separate bivariate tables for the two grade levels for each of the ten achievement

---

2 See Chapter IV for a more detailed explanation of the concept of mental age grade expectancy.

3 A Guidance Handbook for the Adjustment Service in the Elementary Schools, Chicago Public Schools, 1954, 27-28.

tests and has examined these tables in light of the theoretical concept of M.A.G.E. One axis of these tables is used for the data from the Metropolitan Achievement Tests and the other for the California Test of Mental Maturity. Mental ages for the C.T.M.M. are grouped in fifteen half-year intervals and the achievement data are expressed in grade equivalents. Medians, as well as the 25th and 75th percentiles for each of the fifteen mental age intervals were calculated and compared with the expectancy provided by the mental age grade expectancy concept. This type of analysis is essentially what Robert L. Thorndike has described as concurrent prediction.[4] This type of study deals with a number of variables all obtained at approximately the same time, in which one variable has been designated as the criterion and an attempt is made to predict it from the others. Essentially, this is the type of strategy which is operative when an intelligence test and an achievement test or tests are given and we deal with the discrepancies between the two scores. This type of strategy is essentially the same whether we use the continuous distribution of discrepancy scores or we deal with two extreme groups that we label "over achievers" and "under

---

4 Philip A. DuBois and Edward V. Hackett, The Measurement and Evaluation of Over and Underachievement, Contract Nonr (816), Naval Research, Washington University, St. Louis, 83-85.

achievers".

The strategy of concurrent prediction appears to involve a very basic dilemma. This can be stated simply as the problem of differentiating between what is legitimately a prediction and what is really part of the criterion. As an illustration, the typical intelligence test is likely to contain a test of vocabulary and so does the typical test of reading achievement. If the intelligence test is used as a prediction of the reading achievement test, the correlation between the two of them will arise in part at least because they include the same material. Thus, the correlation between the two is only in part a matter of meaningful prediction, and in part a matter of the contamination of one measure by what is inherent in the other.

In an effort to evaluate the extent of this contamination of the C.T.M.M. as a predictor by what is inherent in the ten achievement tests of the Metropolitan battery, a sample of one thousand cases was drawn and a correlation matrix was prepared and studied in light of Thurstone's centroid method of factor analysis. Thus, through this later type of analysis it will be possible to study the extent of the overlap between the C.T.M.M. and the Metropolitan Achievement Battery, and shed some light on the use of the C.T.M.M. as a predictor for the Metropolitan Achievement Battery.

CHAPTER II

REVIEW OF THE RELATED LITERATURE

In the past forty years a variety of methods have been proposed and utilized to relate measures of capacity to measures of achievement. The amount of literature related to this problem is nothing short of voluminous. This chapter presents some of the methods which are most relevant to the problem of relating measures of capacity to measures of achievement.

The methods which have been utilized most frequently are the achievement quotient and accomplishment quotient techniques, and scattergrams. Because the quotient techniques have had such a far reaching effect they will be treated in more detail than scattergrams.

Quotient Techniques of Relating

As William Angoff has pointed out, some of the methods of relating measures of intelligence to measures of achievement which have been developed, are the direct result of the ease of interpretation to which the I.Q. apparently lends itself. In fact, he suggests that the educational quotient and achievement

quotient are the direct result of this type of interpretation.[1]

One of the first to use the term "educational quotient was Raymond Franzen. He defined the educational quotient as "the quotient resulting from the division of the age level reached in the achievement test in question, by the chronological age of the pupil. It is a measure of the rate of progress of the child in the school subject under consideration."[2] His next step in the development of a method of relating achievement to mental ability was a logical one if his first premise is accepted.

> Since the I.Q. is the potential rate of progress and the E.Q. is the actual rate of progress, the ratio of E.Q. to I.Q. gives the percentage of what that child could do, that he has actually done. We can, then, measure the approximation to ideal educational performance of any one child in any one elementary school subject through the approximation of this accomplishment quotient to 1.00. One's differences when E.Q. is subtracted from I.Q. are always positive when they are large enough to be significant and small enough to seem spurious when they are negative.[3]

Independently, at the same time Walter Monroe was developing a somewhat similar concept in relating achievement

---

1 William H. Angoff, Encyclopedia of Educational Research, ed., Chester W. Harris, New York, 1960, 815.

2 Raymond H. Franzen, "The Accomplishment Quotient Technic," Teacher's College Record, New York, XXI, 1920, 435.

3 Ibid., 436.

and intelligence test data for the Illinois Examination. Briefly, his plan consisted of establishing for the achievement tests mental age norms which could be used to supplement the usual grade norms. For each half year of mental age, as shown by the general intelligence scale used, the median achievement was determined. Thus, an achievement quotient of 100 meant that the pupil had achieved exactly the average of the pupils of his mental age or that he was just up to the norm for his mental age. If his achievement quotient was 130, he had achieved thirty per cent more than the average of the pupils of his mental age.[4] It is in this latter point that there was a fundamental difference between Franzen and Monroe's methods. Franzen's method did not tolerate accomplishment quotients over 1.00--they would be "spurious." Whereas Monroe's method not only permitted but expected achievement quotients over 100. In fact, Monroe demonstrated that a distribution of achievement quotients would follow a normal curve.[5]

In spite of what would appear to be a more realistic interpretation on the part of Monroe, it was not his method

---

4 Walter S. Monroe, "The Illinois Examination," Bureau of Educational Research Bulletin No. 6, IX, Oct. 1921, 5-35.

5 Ibid., 56.

which was "picked up" in the literature and practice but rather it was Franzen's method. No doubt this was at least in part due to the fact that Franzen's proposed method would be applicable to any achievement test, whereas Monroe's method could only be utilized with those tests for which mental age norms had been developed.

McCall, in his text How To Measure In Education, was one of the first to advocate the use of the accomplishment quotient.

> The Accomplishment Quotient is the most exact present day measure of the efficiency of study, instruction, and supervision; it is the only just basis for reporting to parents and for judging pupils; and it is the best index of what pupils need special attention and spurring, of what pupils need restraining perhaps, and of what pupils need to be "let alone."[6]

Even more laudatory in their praise of the accomplishment quotient were Stebbins and Pechstein. "We believe that this Accomplishment Quotient is the fairest and most valuable measure now known of the efficiency of the pupil and the teacher."[7] They continue by saying:

> Educational test results taken alone are no fairer to the child than to the teacher. The bright child

---

6 W. A. McCall, How to Measure in Education, New York, 1922, 86.

7 Rena Stebbins and L. A. Pechstein, "Quotients I. E. and A." Journal of Educational Psychology, XIII, Oct. 1922, 392.

> receives the high score and the praise; the duller child takes the low score and defeat, with no regard given to the comparative mentalities. The Accomplishment Quotient is a just measure of the pupil's efficiency in school work.
>
> (1) For the bright children it shows which child is living up to his possibilities and which child fails to make his attainment equal to his capacity to attain.
>
> (2) For the dull children, it shows which child is needing to be urged and helped still more, which child needs restraining, perhaps, and which are most deserving of praise.
>
> (3) Of all children, it asks that the pupil be urged to progress at a rate which is proportional to the mental capacity with which nature endowed him. This is the only fair standard for any child.[8]

Not only was the accomplishment quotient used to evaluate pupil progress in relation to intelligence, but some advocates even suggested that teacher competency could be measured by it. Thus, it seemed a long sought after educational tool had been uncovered.[9] The use and application of A.Q. became quite widespread and it was in this use that even more flagrant misinterpretation developed.

In a study conducted by Pintner, he discovered that not only were pupils of below average intelligence doing

---

8 Ibid., 395.

9 This position was taken by Stebbins and Pechstein, see Ibid., 392, as well as by I. N. Madsen, "Interpreting Achievement in School in Terms of Intelligence," American School Board Journal, Nov., 1921, 41 and 111.

better than expected in achievement tests, but also conversely the bright pupils as a group were not achieving up to expectations.[10] This led him to conclude:

> We have been pushing and cramming the duller children, while the bright ones have been allowed to loaf. The bright child is the most retarded child in our schools. The dull child is the most accelerated. The bright child is the laziest child and the dull child is the most industrious....Habits of mental laziness acquired in school often persist through life, and there are undoubtedly many adults at the present time, who have splendid native ability and who do not know it, because the school has taught them to be satisfied with a mediocre type of accomplishment.[11]

The studies of McPhail,[12] Murdock,[13] as well as the study by Beeson and Tope[14] seemed to confirm the discovery made by Pintner. That is, the brighter pupils were not achieving in relation to their mental ability, as well as the below average pupils were achieving in relation to their mental ability.

---

10 Rudolf Pintner, "The Significance of Intelligence Testing in the Elementary School," The Twenty-first Yearbook of The National Society For The Study of Education, Bloomington, 1922, 166.

11 Ibid., 167.

12 Andrew McPhail, "The Correlation Between I.Q. and A.Q." School and Society, XVI, Nov. 1922, 586-88.

13 Murdock, "The A.Q., Finding and Using It," Teachers' College Record, XXIII, May, 1922, 229-39.

14 M. F. Beeson and R. E. Tope, "E.Q. and A.Q. as Aids in Classification of Pupils," Journal of Educational Research, IX, April, 1924, 281-92.

Once the "fact" was accepted that the bright students were not achieving as well as might be expected, still other research workers attempted to uncover the reasons why this was so. Torgerson and Shuman, as a result of a study of 216 pupils, concluded the reason the accomplishment quotient bears an inverse ratio to the intelligence quotients is not because this relationship is an inherent one, but rather that it has been produced by the traditional school with its artificial and improper grading system.[15] Others felt that once teachers, pupils and parents were appraised of their ability there would be a consequent change due to more effort.

One of the first to question the thinking behind the accomplishment quotient was J. Crosby Chapman, who also pointed up one of the reasons why such a technique was so appealing.[16]

> With the entrance of intelligence tests and school tests, it was a great temptation to measure intelligence and school achievement, and then by the difference in standing to estimate the extent to which an individual was taking advantage of his school opportunity. The general idea is so attractive and the results, if true, so useful that school men have been captivated by the simplicity of a definite figure which

---

15 T. L. Torgerson and Irene Schuman, "The Variability of the Accomplishments of Pupils of the Same Mental Level," Journal of Educational Research, XI, Feb. 1925, 133-134.

16 J. Crosby Chapman, "The Unreliability of the Difference Between Intelligence and Educational Ratings," Journal of Educational Psychology, XIV, Feb. 1923, 103-108.

promised to give such valuable information with regard to the pupil and the school.[17]

Chapman's investigation from two separate sets of data on intelligence tests and achievement tests led him to conclude that in order to predict a differential educational index with any degree of accuracy (.93 correlation), it would be necessary to repeat the test at least six times.[18]

Even more critical of the accomplishment quotient technique was Herbert Popenoe, a statistician for the Los Angeles City Schools. In an effort to determine the probable error of the accomplishment quotient, an extensive study was undertaken dealing with more than six hundred pupils ranging from third to eighth grade. Each student took two forms of an intelligence test, an arithmetic test, reading test and a spelling test. As a result of the two forms of the tests it was possible to calculate two comparable accomplishment quotients for each pupil. A correlation of these two accomplishment quotients was then calculated yielding a correlation of .28 which in effect was the reliability of the accomplishment quotient. Additional studies were made of the correlation between the accomplishment quotient and the intelligence quo-

---

17 Ibid., 103.

18 Ibid., 108.

tient for groups of individual students, as well as a correlation between the mean accomplishment quotient and the mean intelligence quotient, for twenty-four schools. These correlations were -.39 and -.59 respectively. These studies led Mr. Popenoe to conclude:

> This reliability (of the accomplishment quotient) is lower than the minimum desirable to continue a statistical measure in use . . . pupils do not have an equal opportunity to attain high accomplishment quotients. Schools have even less equal opportunity to attain high mean accomplishment quotients, than have pupils to attain high individual accomplishment quotients. Insofar as may be judged from the results of this study, the administrative use of the accomplishment quotient is open to serious criticism.[19]

T. L. Kelley in his book Interpretation of Educational Measurement took a similar position as to the weaknesses of the accomplishment quotient. "The only sound way to judge the efficiency of a particular kind of accomplishment quotient is to check up on all cases, average as well as extreme." Such an examination, the author predicted, would lead to the conclusion that in the majority of cases the quotient would be found to be unreliable.[20] Even more shattering and to the core

---

19 Herbert Popenoe, "A Report of Certain Significant Deficiencies of The Accomplishment Quotient," Journal of Educational Research, XVI, June, 1927, 46-47. Simultaneous with the publication of this study the Los Angeles city schools discontinued the use of the accomplishment quotient.

20 Truman Lee Kelley, Interpretation of Educational Measurements, World Book Company, 1927, 24-25.

of the nature of the problem caused by the accomplishment quotient was what he called the "jangle fallacy," which he defined as ". . . the use of two separate words or expressions covering in fact the same basic situation but sounding different, as though they were in truth different."[21] Achievement and intelligence sound as though they are different; they have different jangles but in truth are they different? In an effort to test this difference as to whether intelligence test results are different from achievement test results, Kelley undertook to find what part of the variance in the two tests would be common to both and what part would be different. ". . . we would conclude that ninety per cent of the two traits (achievement and intelligence) correlated was identical and ten per cent was different."[22] Thus Kelley, as early as 1927, demonstrated that the study of differences between achievement tests and intelligence tests were based upon ten per cent of what was being measured.

William Wilson of Ohio State University took up the question of the accomplishment quotient and came to the conclusion that at the least it was misleading. In answer to the issue that bright students have low accomplishments and dull students have high accomplishment quotients, he said that

---

21 Ibid., 64.

22 Ibid., 195-196.

even by an inordinate effort the bright student cannot raise his accomplishment quotient to unity, and that, even if every dull student were as lazy as the bright are commonly believed to be, the accomplishment quotients of this group would still be gratifyingly large.[23] This conclusion was reached and influenced by a study of the correlations of accomplishment quotients and I.Q.'s as well as a subjective recognition of the fact that achievement is not solely influenced by ability, but is also influenced by a myriad of factors, "health, home surroundings, sensory acuity, nutrition, and personality."[24]

In spite of the work done by the critics of the accomplishment quotient, the accomplishment quotient continued to be used. In 1928, P. H. Nygaard, after recognizing the validity of the work done by Wilson, offered a formula which would eliminate negative correlations between accomplishment and I.Q. by cancelling out the influence of regression.[25]

Genevieve L. Cory of Teachers College, obviously unaware of the work done by the critics of the accomplishment

---

23 William Wilson, "The Misleading Accomplishment Quotient," Journal of Educational Research, VII, January, 1928, 2-3.

24 Ibid., 9.

25 P. H. Mybaard, "A Revised Accomplishment Quotient," Journal of Educational Research, XVIII, June, 1928, 87.

quotient, reported on a study made of the changes in the accomplishment quotients for pupils of grades 3A to 6A in the public schools of Shaker Heights, Ohio. After examining the various changes, she concluded teacher efficiency could be evaluated in terms of the accomplishment quotients of their pupils, provided the teachers have begun the year with comparable groups, and when the differences found can claim high statistical reliability. This reasoning was based upon the fact that she found pupils who had high I.Q.'s (130 and above) and low A.Q. showed greater gains in terms of A.Q. than did pupils with I.Q.'s of 81 to 90. In fact she found the A.Q.'s of this latter group went down.[26]

Clyde Morley of the University of Wisconsin was also concerned about the use and application of measuring pupil accomplishment by such a simple and plausible device as the achievement quotient which he defined as ". . . dividing the subject age obtained . . . of the achievement test by the mental age of the intelligence test." It is unfortunately confusing that Morley used the term achievement quotient since this is not the same concept that Monroe was referring to, but rather is the

---

26 Genevieve L. Cory, "A Study of Various Factors Which Influence the Use of the Accomplishment Quotient As a Measure of Teaching Efficiency," Journal of Educational Research, XXI, Jan. 1930, 29-42.

same term as developed by Franzen. Morley's calculation was more direct in that he bypassed the need to cancel the chronological ages which is necessary in the $\frac{\text{"E.Q."}}{\text{I.Q.}}$ formula. His study which was based on sixth grade pupils in the Racine Public Schools was concerned with several aspects of the problem created by the practice of comparing mental ability test data and achievement test data, namely: the influence of mental test reliability, the influence of achievement test reliability, the influence of the intercorrelation of mental and subject tests, the possible effect of pushing pupils and the use of achievement quotients for individual diagnosis.[27] Although his conclusions were not too different from the other critics of the accomplishment quotient, he did point out two very important points which had been overlooked by the "followers of Franzen" that is,

> The use of the achievement quotient technique involves the assumption that accomplishment, in a particular school subject and intelligence are distinctly separate traits. . . . Achievement quotients sufficiently reliable for individual diagnosis cannot be derived from a single administration of present group tests, except for subjects having a low correlation with intelligence.[28]

---

27 Clyde Morley, "The Reliability of the Achievement Quotient," The Journal of Educational Psychology, XXI, May 1930, 351-360.

28 Ibid., 360.

Although the Thirty-Fourth Yearbook of the National Society for the Study of Education which was devoted to Educational Diagnosis did not so much as even refer to A.Q., it was not so easily set aside. It reappeared in the literature and in some instances still has some supporters. In fact, Leonora E. Loughin justified its use in evaluating the achievement of pupils in a comparative study of two public schools as late as 1937.[29] One might justifiably ask how it is possible that after the many studies which have been cited, that someone would utilize this questionable practice. It is here that we have what might be considered the "research lag."

Generally, authors of textbooks which are concerned with a wide range of topics find it extremely difficult if not impossible to include the latest research on all topics included in their texts. In part this is due to the volume of research which is being conducted and in part to the length of time which it takes to get a proposed book from manuscript to salable copies. Coupled with this "problem" is the difficulty that confronts the practitioner when he attempts to utilize workable and proven techniques. If he spends all of his time determining what is the most feasible method to be

---

29 Leonora E. Loughin, The Relative Achievement of Pupils of Two Public Schools, M.A. Thesis, University of Chicago, 1937.

utilized as demonstrated by research, he does not have sufficient time to put these ideas into operation. These time factors are further compounded by the problem that is created by the fact that once an idea is put into print it seldom is changed or marked: incorrect, outdated, or superceded by, some other research. There is still another plausible factor which contributes to this problem. This is the situation which develops as the result of having been taught an accepted practice which has later been found to be untenable in light of more recent research.

In studying the Accomplishment Quotient, this writer found particularly good evidence to demonstrate the latter aspects of this problem. In their text, The Use and Interpretation of Elementary School Tests, which was published in 1937, Greene and Jorgensen refer to the Accomplishment Quotient or Accomplishment Ratio as "nothing more than the ratio between the relative educational development of a pupil (E.Q.) to his relative brightness (I.Q.)." According to them, teachers can utilize this approach to determine whether a certain pupil is making progress in his school work that he is capable of making.[30] Although this text was published in 1937, no mention was made of

---

30 Harry A. Greene and Albert N. Jorgenson, The Use and Interpretation of Elementary School Tests, New York, 1937, 242-244.

any of the research which would indicate that there were some serious statistical questions which had been raised to question this concept. This text, which was "designed especially for the elementary school teacher and the student of elementary education," was in all probability quite widely used since there were relatively few others available.[31] What is an even more interesting speculation and beyond the scope of this paper is, do any students of elementary education still use this text?

A further illustration of this same research lag is seen in a text Practical Measurements for School Administrators, designed "for students of education and as a handbook in measurements for the school administrator."[32] According to Noble "intelligence tests may be employed to measure capacity to learn and school achievement tests to measure school achievement." He then goes on to define various equations of real significance.

$$EQ = \frac{E\ A}{C\ A} \times 100$$

The educational quotient equals the educational age divided by the chronological age multiplied by 100. The

---

31 Ibid., V. The 1954 high school revision of this text is critical of the A.Q. but claims it is applicable to groups.

32 M. C. S. Noble, Practical Measurements for School Administrators, Scranton, Pennsylvania, 1939. VII.

achievement or accomplishment quotient is obtained by utilizing either of two equations which are:

$$(a) \quad A\,Q = \frac{E\,A}{I\,Q} \times 100$$

$$(b) \quad A\,Q = \frac{E\,A}{M\,A} \times 100$$

"Of the two equations for measuring the pupil's A.Q., the second equation is obtained more simply, hence is used most frequently. It should be understood, however, that although it is possible to calculate the A.Q. of the average child in a group, the A.Q. is fundamentally a measure to be used in the study of an individual pupil."[33] Unless Noble was trying to introduce a new variation of the accomplishment quotient his equation (a) included a mistake or at least a printer's error. If E.Q. were substituted for E.A. in (a) then the two equations would be comparable for:

$$A\,Q = \frac{E\,Q}{I\,Q} \times 100 = \frac{\frac{E\,A}{C\,A}}{\frac{M\,A}{C\,A}} \times 100$$

Not only was this error included, but once again no reference was made to any of the previously cited studies which would

---

33 Ibid., 270-271.

raise doubts in the reader's mind as to the possible limitations of this method.

The recurring nature of the problem created by the development and use of the accomplishment quotient is further testified to by the fact that as recent as 1962 William Asher published an article to demonstrate the erroneous conclusions which can and have been drawn by the use of the accomplishment quotient.[34]

Although the accomplishment quotient has had a tendency to recur, it has gradually given way to the "Age of the Scatter Diagram" as a method of relating measures of intelligence to measures of achievement.[35]

Scatter Diagram Technique

Scatter diagrams are variously referred to in the literature as correlation tables, two-way tables, bi-variate tables, double-entry tables, comparison charts, quintile classification charts, scattergrams and scatter diagrams. Essentially, all are similar in that they are graphic methods

---

34 William Asher, "Statistical Problems of the Accomplishment Quotient," Journal of Experimental Education, XXX, March 1962, 285-287.

35 David Tiedeman and Charles C. McArthur, "Over and Underachievement: If Any," 13th Yearbook National Council of Measurements Used in Education, New York, 1956, 135-45.

of showing the relationship between two sets of variables, where one axis of the diagram represents one variable and the other axis of the diagram represents another variable. Broadly speaking, they may be divided into two groups, the first four terms usually referring to charts designed to assist in the calculation of the correlation between two variables, the remaining terms usually referring to charts designed to reveal pertinent information obtained by comparing two variables for each individual in a group. Where the prime purpose of the scattergram is to show the relationship between two scores or combination of scores for each individual, it is evident that the pairs of scores must represent the same individual. Perhaps the best definition of the charts under consideration is given by Goodenough: ". . . the figure resulting when scores in one of two related variables are plotted along the abscissa of a curve, with those of the others plotted along the ordinate, the number of cases at each level being recorded at the appropriate points of intersection between the rows and columns."[36]

Charles Germane and Edith Germane in their text, Personnel Work in High School, outline six steps to be followed in the construction of a quintile scattergram.

---

36 Florence Goodenough, Mental Testing, New York, 1949, 565.

(1) Rank pupils in descending order of merit based on the results of an "ability to learn" test.
(2) Rank these same pupils in descending order based on a standardized test in the subject.
(3) Make five equal, or nearly equal, divisions in each of the two columns.
(4) Draw a five-inch square and divide it into 25 one-inch squares. Number these small squares from left to right across the large square. Number the initial row of small squares: I, II, III, IV, V.
(5) From top to bottom, number the rows of small squares: 5, 4, 3, 2, 1.
(6) Transfer the Quintile Ranking to the Quintile Chart by tallying each individual according to his ranking on the two tests.[37]

Some of the advocates of the scatter diagram as a method of relating test data, also suggest taking an additional step in the preparation of such a chart. They suggest dividing such charts into four sections or quadrants and the pupils of each of these quadrants can be tentatively classified in four groups:

(1) Upper right quadrant: high achievement, high ability
(2) Lower right quadrant: low achievement, high ability (under-achievers)
(3) Upper left quadrant: high achievement, low ability (over-achievers)
(4) Lower left quadrant: low achievement, low ability.[38]

---

37 Charles Germane and Edith Germane, Personnel Work in High School. New York, 1941, 97-115.

38 D. W. McDonald, "Scatter Diagrams and Their Use," The School Guidance Worker, XVI, March, 1961, 5.

Pressey states that when the teacher gets the results of tests before her in tabular form instead of stopping with unwieldly class lists, at least half of her troubles in handling and interpreting the data will be eliminated. He goes on to claim that perhaps no single method will be found more useful by the teacher for recording sets of scores than the scatter diagram.[39]

The advantages of the scatter diagram as a means of relating test data according to its advocates are:

(1) It discloses which pupils are working up to capacity.
(2) It is a method of analyzing the extent to which the group as a whole is working at, above or below capacity.

Bowman offers an even more sophisticated use of the scatter diagram. He advocates a third dimension to the charts by using different colors for pupils of a particular interest, age, marks in school or other pertinent information. He also is quick to point out one of the inherent weaknesses of the scatter diagram as follows: "Obviously the fact that the pupil is indicated on the chart as a deviate does not mean that this is his true status. This could only be true if the tests used

---

39 Sidney L. Pressey, and Luella Cole Pressey, Introduction to the Use of Standardized Tests, New York, 1923, 40.

were completely reliable."[40]

As will be seen in the following chapter, there are several school systems that use such a technique quite extensively. In spite of this rather wide usage, a survey of the literature revealed but a few sources of information regarding their use and development.

Some Recent Techniques

In the past ten years the effort to uncover a method of relating mental ability test data to achievement test data has continued. Two more recent techniques to be suggested are Durost's Modal Age Norms and Shanner's Anticipated Achievement. Both of these techniques attempt to overcome some of the inherent weaknesses of the previously mentioned techniques.

Durost would have the test publishers develop modal age norms for both the capacity and achievement measures.[41]

---

40 Howard A. Bowman, Techniques For Graphical Representation of Pupil Personnel Data to Indicate Individual Deviates and to Provide a Basis for More Adequate Guidance, Educational and Psychological Measurement, XII, No. 3, 1952, 390-391.

41 Modal age norms are grade equivalents or other devised norms that are computed for those pupils who are at the proper grade level for their age, rather than for all pupils in the grade level. Pupils who are retarded or accelerated are not included in the norm group. The standardization groups used for modal age norms are found to be slightly superior as compared to a complete norm group. T. L. Kelley was one of the first to advocate the use of such norms.

This he feels would insure the comparing of youngsters who for all practical purposes had progressed through school at a normal rate. He would then have the publisher develop deviation scores for each of the two different types of tests so that a study of any deviations for a group or student could be made.[42]

Shanner's approach to the problem of relating mental ability to achievement is somewhat similar to the proposal made by Durost. Shanner also recognized the problems created when tests which have been standardized on uncontrolled populations are compared. The result of such standardization procedures is the sample for norming is homogeneous with respect to one variable, usually grade for achievement tests, but which is rather heterogeneous with respect to variables such as age and mental ability. His solution to this problem was to calculate the mean achievement scores of a large number of nationwide samples of pupils homogeneous with respect to grade, age and mental ability. Thus when the grade, age and mental ability characteristics of a pupil are known, one would merely compare such a pupil with the sampling group most nearly the same grade,

---

42 W. N. Durost and G. A. Prescott, "An Improved Method of Comparing a Capacity Measure and An Achievement Measure At the Elementary School Level," Educational Psychological Measurement, XII, 1952, 741-751.

age and mental ability.[43]

Summary

In the past forty years, a variety of methods and techniques have been offered in an attempt to relate measures of intelligence to measures of achievement. Some of these techniques have been examined in this chapter. Particular attention has been given to the accomplishment quotient technique, since it of all techniques has been the one which has had the greatest number of followers and has been the most difficult to eradicate in spite of its inherently recognized shortcomings.

The subsequent chapter will present some of the practices which are currently being utilized by large public school systems.

---

43 William Shanner, "New Concepts In Norms," Positive Values in the American Educational System, American Council on Education, Washington, D.C., 1959, 64-74.

CHAPTER III

CURRENT PRACTICES AS REVEALED BY A QUESTIONNAIRE

In an effort to collect information regarding current practices of relating mental ability test data to achievement test data, a questionnaire was designed and sent to fifty-six cities with public school enrollments of two-hundred thousand pupils and over.[1] The cover letter which accompanied the questionnaire was directed to the superintendent of schools of each of the cities. This letter described the purpose of the questionnaire and sought their assistance in having it directed to the person in their school system best qualified to answer questions regarding their testing program.

Within three weeks of the first mailing twenty-nine, or 51 per cent of the fifty-six had responded. A second mailing with a duplicate questionnaire and cover letter was sent to the twenty-seven who had not responded. Four weeks after this second mailing, an additional twenty cities responded for a

---

1 The list of cities and a copy of the questionnaire may be found in Appendix I.

total of forty-nine respondents or 87 per cent of the forty-six cities which had been selected.

Of the forty-nine questionnaires returned forty-eight of them were usable. One questionnaire was returned without being answered due to a change in administration. The subsequent part of this chapter analyzes the responses of the forty-eight cities which responded to the questionnaire.

The first question of the questionnaire asked, "Do you have a city-wide standardized testing program?" Forty-seven of the forty-eight cities indicated that they did have a city-wide testing program. The only city to respond negatively to this question was Jersey City. Thus, ninety-seven percent of those responding indicated that they had a city-wide standardized testing program. An examination of the kind of program was possible through an inspection of the responses to the second part of the first question, which asked for a description of the testing program in terms of the grade levels tested as well as the mental ability tests and achievement tests used.

In examining the responses to the latter part of the first question, it becomes quite apparent that there was a great deal of variation not only in the grades tested but also in the tests used. Some cities tested as frequently as every year while others tested only once in elementary school and again in high school. Further examination of the responses to

this part of the questionnaire indicated that the intelligence tests most frequently used were the Kuhlmann Anderson, California Test of Mental Maturity, and Otis, although other tests such as the Hennon Nelson, S.R.A. Primary Mental Abilities and Pintner Cunningham were being used by a few cities. The achievement tests most frequently utilized were the Metropolitan, Stanford and the California tests, although once again there were others being used such as the Iowa Tests of Basic Skills and S.R.A. Achievement Series.

Although the tests used and the grade levels tested were of interest, this part of the questionnaire was designed to determine the extent to which intelligence tests and achievement tests were being administered at the same grade level. An analysis of the extent to which such a practice is current in grades one to eight is revealed in the following table.

TABLE I

NUMBER OF GRADE LEVELS AT WHICH BOTH INTELLIGENCE AND ACHIEVEMENT TESTS ARE ADMINISTERED IN GRADES 1 TO 8

No. of Grade Levels	No. of Cities	Per cent of Cities
5	3	6
4	4	8
3	11	23
2	19	38
1	6	13
0	5	11

Inspection of Table I reveals the rather common practice of administering an intelligence test as well as an achievement test at the same grade level. In fact, out of the forty-eight cities no less than forty-three of them did this at one grade level or more. Interestingly enough eighteen cities did this as frequently as three or more times within the period from first to eighth grade.

The second question asked for information regarding methods which the various cities used to relate mental ability test results to achievement test results for individual students. An analysis of the responses to this question indicated some very divergent thinking as to the methods which should or should not be used. Table II presents a summary of the responses to this question.

TABLE II

METHOD UTILIZED TO RELATE MENTAL AND ACHIEVEMENT TEST DATA

Method Used	No. of Cities
None	14
Scattergrams	8
Stanine	6
Expectancy Tables	6
Teacher	3
Principal	2
Percentile Rank	2
Profile	2

TABLE II (continued)

METHOD UTILIZED TO RELATE MENTAL AND ACHIEVEMENT TEST DATA

Method Used	No. of Cities
Anticipated Achievement	2
Intelligence Grade Placement	1
Achievement Index	1
Publisher's Material	1

A more detailed examination of the methods which were used by four or more cities was made. This examination included a study of the following methods: scattergrams, stanine and expectancy tables.

Scattergrams

Scattergrams, as a method of relating mental ability test results to achievement test results for individual students, are used by eight cities: Miami, Oakland, Omaha, Rochester, Toledo, Wichita, Phoenix and Tulsa. Miami, Oakland and Wichita provided test manuals of their procedures while the others indicated this method was used but did not provide descriptive literature.

The Miami school system uses scattergrams based on mental ages and achievement age equivalents. The age equiva-

lent scores are derived from a conversion table provided by the city. In their manual describing this procedure they point out that due to the lack of complete reliability of tests, a span of one year is provided in determining whether a pupil is working at his capacity level. Furthermore, they point out that scattergrams indicate discrepancies based on just one achievement test and one mental ability test and they do not indicate why these discrepancies exist.[2]

Wichita also uses group analysis charts (scattergrams) to relate mental ability test scores to achievement test results. Rather than using mental ages or grade equivalents, they convert the intelligence quotient as well as the achievement test results to percentiles and base their scattergrams on these. The scattergrams which they have developed have a band score. Thus a pupil scoring at the fiftieth percentile on a mental ability test would be expected to score somewhere between the seventy-fifth and twenty-fifth percentiles on the achievement test. These latter two points are referred to as the approximate boundaries of normal relationship. Cases which would deviate from this normal pattern would undergo further investigation.[3]

---

2 The County-Wide Group Testing Program, Dade County, Florida Public Schools, 106-107.

3 Test Handbook for the Elementary Schools, Research Department, Wichita Public Schools, 1962, 33-34.

Scattergrams are also used in Oakland. They refer to these as Class Study Charts. Their method is based on a system of comparing the mental age and achievement ages for an individual with the median mental ages and achievement ages for the total grade level. A table is provided indicating the expected median mental and achievement ages for second grade through ninth grade. Once a given pupil's mental age and achievement age have been compared to the expected for that grade level he is assigned two letter classifications, one for his mental age and one for his achievement age. These letter classifications, ten in all, are based on a six-month deviation scale from the expected median.

These letter designations are then plotted on a two-way chart which has ten classifications for achievement and ten for mental ability. The basic purpose of these charts as indicated by Oakland "is to assist the teacher to obtain a better picture of his pupils in terms of their mental and achievement levels." It would appear that this method of plotting a scattergram would be somewhat more difficult since both the mental age and achievement age would have to be derived by a study of the deviation from the median.

Expectancy Tables

Six cities: Denver, Los Angeles, Milwaukee, New York,

St. Paul and Washington, D.C. indicated that they used expectancy tables to relate achievement test results to mental ability test results for individual students. Washington, D. C., and St. Paul both indicated they use the expectancy tables which are provided by the test publishers, however they did not indicate which test publisher. Milwaukee merely indicated expectancy tables were utilized but did not detail their method.

New York, Denver and Los Angeles did provide descriptive materials which explained their techniques. New York uses an expectancy score formula which is:

$$\frac{\text{I.Q.}}{100} \times \text{Grade Equivalent} = \text{Expected Grade Equivalent}$$

Thus a pupil with an I.Q. of 150 in the fourth month of sixth grade would have an expected grade equivalent score of 9.6. Denver and Los Angeles have developed much more detailed types of expectancy tables.

Denver has developed Expected Achievement Grade Tables which are based on formulas which weight the mental age. The formulas that were used to determine the achievement ages were as follows:

Age	
6 yrs. to 8 yrs. 6 mos.	$\frac{MA + CA}{2}$

8 yrs. 6 mos. to 10 yrs.	$\frac{3\ MA + 2\ CA}{5}$
10 yrs. to 12 yrs.	$\frac{2\ MA + CA}{3}$
12 yrs. and above	$\frac{3\ MA + CA}{4}$

From these formulas it was possible to develop for each grade level a table indicating the I.Q. and the expected grade placement achievement for each of the ten school months.

These tables permit a teacher not only to determine the grade placement achievement for an individual pupil, but it is also possible for a teacher to use these same tables to determine the increment of growth or progress which may be expected of a pupil of a given grade and I.Q. An examination of these tables would indicate that a fifth grade pupil of 115 I.Q. may be expected to progress five months between September and February, while a fifth grade pupil with an I.Q. of 90 may be expected to progress three months in the same period.

Denver feels that the value of the expected achievement tables as used by them is realized through establishing understandable attainable goals for pupils.[4]

---

4 Expected Achievement Grade Placement Tables, Division of Instructional Services, Denver Public Schools, 1962, 2-3.

The Los Angeles Expected Achievement Grade Placement Tables are derived from a series of formulas identical to the ones used by Denver.[5] However, instead of expressing these tables by grade levels as is done in Denver, they are done by chronological age from six years to sixteen years six months. They also differ insofar as expectancies are derived for I.Q.'s from 60-140, whereas the Denver tables only included the group of 70-130. Howard Bowman, who is the Los Angeles Director of Evaluation and Research, in a supplemental statement indicated that recent studies have indicated that the expectancy concept being used is not sufficiently reliable for other than group use and currently more comparisons with grade norm are being used.

Stanines

There were six cities which indicated they used stanines to relate mental ability test results to achievement test results for individual students. These six were Tampa, Baltimore, Tucson, San Diego, Jacksonville and Kansas City. Of these six, four of them provided descriptive literature of their techniques. Baltimore and Tampa indicated they used stanines

---

5 Expected Achievement Grade Placement Tables, Division of Instructional Services, Los Angeles City School Districts, 1955, 6.

but did not provide explanatory data.

Stanines are a type of standard scores which have a mean of five and a standard deviation of two. The units of the stanine scale range from a low of one to a high of nine. Stanines, like other standard scores, have the advantage that each stanine represents the same spread in ability. Hence the difference of stanine one and two is the same as the difference between stanine eight and nine.

Tucson, Arizona in their test manual cite some of the advantages in using stanines: stanines give a far more accurate index to a student's probable ability or achievement than either a precise appearing I.Q. or a grade equivalent; stanines reduce all test scores to comparable units subject to comparison; stanines give an accurate picture of the range of abilities within a group; stanines are the only acceptable terms for interpreting test results to parents and students and stanines greatly facilitate recording, reporting and I.B.M. data processing.[6] Even though stanines are band scores the Tucson public schools are urged to be cautious in interpreting small differences in mental ability and achievement. They recommend that a difference of one stanine score be regarded as a chance

---

6 Group Testing Bulletin II Testing Office, Department of Guidance and Counseling, Tucson Public Schools, 1962, 22-23.

factor. A difference of two stanines is considered significant; a difference of three stanines is a very real difference and something more than chance is operative.[7]

San Diego is also careful to see that the difference in mental ability stanine and achievement stanine is not misinterpreted. They accomplished this by using a profile card which indicates the stanine ability level by a band one stanine above and one stanine below the mental ability stanine for an individual. "The three-level width of the band is because of the inadequacy of any test of intelligence to measure ability level within narrow limits."[8]

Kansas City and Jacksonville did not explain the use of stanines in their cities other than to indicate that all tests are converted to such a scale for purposes of relating mental ability and achievement data.

Other Techniques

The techniques most frequently used to relate mental ability tests and achievement test results for individual students as shown by Table II were scattergrams, stanine and ex-

---

7 Ibid., 37.

8 Statement On Profile Of Test Performance Cards, Instructional Division, San Diego City Schools, 1960, 2.

pectancy tables. The other techniques mentioned in Table II in most instances were not described sufficiently to make it possible to report on them. Two cities, Philadelphia and Pittsburg responded quite differently to this question and their answers do deserve some consideration. Both of these are quoted for the reader's information.

Pittsburg in response to this question said:

> We believe spurious interpretations are obtained when individual mental ability test scores are related to achievement test scores. However, comparing the achievement and mental ability of groups has some meaning. Pittsburg schools have never made this comparison in a formal published report.

Philadelphia responded in a somewhat similar manner.

> This is not done. In the final analysis both kinds of tests measure what has been learned. We use achievement tests, in preference to I.Q. tests, as a prognostic device.

The last two questions of the questionnaire requested information about the provisions to insure uniform analysis and the availability of materials describing the techniques of relating mental ability test results to achievement test results. Question three requested information about the practices used by them to insure uniform analysis. Eight cities indicated they did not take any steps to insure such an analysis. Of the remaining forty cities four techniques were identified as being used: meetings, eight; in-service training, four; distribu-

tion of materials, seven; and central office communications, twenty-one.

In response to the question asking for information regarding duplicated materials for their method of relating mental ability test data and achievement test data, twenty-six cities indicated they had developed such materials. Of these twenty-six, fifteen of them sent copies in response to the request for such materials. An analysis of these materials indicated that these materials had been developed as test booklets to help teachers and administrators to carry out their responsibility for the testing program and not specifically for the purpose of relating mental ability test data and achievement test data.

Summary

A questionnaire was sent to fifty-six cities in an effort to collect information about current practices of relating mental ability test data to achievement test data. Forty-nine cities responded and of these responses forty-eight were usable. The practice of administering mental ability and achievement tests at the same grade level was quite apparent. No less than forty-three cities did this at least once in grades one to eight. An examination of the practices of relating mental ability test data and achievement data indicated

that the most frequently used methods were scattergrams, stanines and expectancy tables.

CHAPTER IV

AN EXAMINATION OF A PRACTICE

In the Chicago Public Schools the practice of relating intelligence test results and achievement test results has been accomplished through the use of the Mental Age Grade Expectancy concept. This concept owes its origin to Dr. Grace Munson, who at one time directed the Bureau of Child Study.[1] As early as 1920, when she was working as a psychologist, she used the term in her case study reports. In 1928 when she was in charge of the Guidance Center at Sullivan Junior High School, she used the concept of mental age grade expectancy as one of the factors for developing homogeneous groups. She had additional opportunities to experiment with the idea when she served as the principal at the Kilmer and Ray Elementary Schools. In 1935 when she became the Director of the Bureau of Child Study, she incorporated this concept into a handbook to be used by the psychologists in writing their case studies.

---

1 The writer has corresponded with Dr. Munson who was gracious enough to provide much of the information about the origin of M.A.G.E.

When a handbook for elementary adjustment teachers was developed in 1954, this concept was included as the method to be used to relate mental age and achievement tests. Briefly this concept reduces mental age scores and achievement test scores to an equivalent basis through the use of a mental age expectancy table. The following has been abstracted from the table which appears in the adjustment teachers handbook.[2]

Mental Age			Achievement
Years		Months	Grade Expectancy
6	-	0	1.0
6	-	1	1.1
6	-	2	1.2
6	-	3	1.3
6	-	4	1.3
6	-	5	1.4
6	-	6	1.5
6	-	7	1.6
6	-	8	1.7
6	-	9	1.8
6	-	10	1.8
6	-	11	1.9
7	-	0	2.0

An examination of this table reveals one of the inherent difficulties of relating achievement which is expressed in grade equivalents and mental age expressed in years and months. The educational year has ten months as indicated by

---

2 A Guidance Handbook for the Adjustment Service in the Elementary Schools, Chicago Public Schools, Chicago 1954, 28.

grade expectancy in contrast to mental age which has a twelve month year. In order to adjust this difference it is necessary to repeat two grade expectancy months, 1.3 and 1.8. A closer examination of this table would also lead the reader to conclude that there is a one-to-one relation between achievement and mental age. Furthermore, no effort is made to suggest that deviations from mental age in terms of achievement would vary depending on the type of achievement being measured. Interestingly enough, the handbook for psychologists which was revised the same year the adjustment handbook was published, 1954, incorporated the following statement: "Omit deviation for arithmetic, since achievement in this area is more closely related to grade placement than to mental age."[3] Although no statement is made regarding the significance of deviations the mere word deviation would appear to indicate that something was not following an expected pattern. The question arises, what should be the expected pattern? Are there possibly other deviations which result from tests which will not follow the mental age grade expectancy because like arithmetic they are influenced by such factors as grade placement, chronological age, and still

---

3 Bureau of Child Study, Psychologists' Handbook, Chicago Public Schools, 1954, p. 9.

others.

In order to examine this practice of relating test results,a study was made of the intelligence and achievement tests administered to pupils who were in 6A and 8B during the Spring of 1961 as part of the Chicago city-wide testing program. The 6A pupils were administered the California Test of Mental Maturity Short Form, Elementary Level and the Metropolitan Achievement Battery, Intermediate Level. The California Test of Mental Maturity Short Form, Junior High Level and the Metropolitan Achievement Battery, Advanced Level were given to the 8B pupils.

The California Test of Mental Maturity consists of seven sub-tests which sample various kinds of mental processes to establish the level and rate of mental development. The seven sub-tests are grouped to yield three scores: Language, Non-Language and Total. The sub-tests which contribute to the Language score are designed to measure the rate of development of mental ability as evidenced by responses to verbal situations. In contrast, the sub-tests which contribute to the Non-Language score are intended to measure the rate of development of mental ability as evidenced by responses to non-verbal situations, by which is meant type of material rather than the method of administration. The Total score is an average of Language and

Non-Language scores.[4]

The Metropolitan Achievement Battery comprises a co-ordinated series of measures of achievement in the important skill and content areas of the elementary and junior high school curriculum. The complete battery as was administered consists of ten tests.[5]

Test 1 - Word knowledge is a vocabulary test designed to measure a pupil's knowledge of the literal meaning of words.

Test 2 - Reading is designed to measure various aspects of reading comprehension. This is accomplished by providing a series of reading selections, each followed by several questions.

Test 3 - Spelling is an item recognition type of test whereby the pupil is to judge whether a word is spelled correctly or incorrectly.

Test 4 - Language is a measure of the relative strengths and weaknesses of the language skills appropriate for a particular grade.

Test 5 - Language Study Skills is designed to measure the pupil's ability to indicate the best reference source and the ability to use the dictionary.

---

4 Elizabeth T. Sullivan, Willis W. Clark and Ernest W. Tiegs, Manual: California Short-Form Test of Mental Maturity, Los Angeles; California Test Bureau, 1957, p. 3.

5 Walter N. Durost et al., Manual: Metropolitan Achievement Tests, Chicago and New York, World Book Co., 1959, p. 3-4.

Test 6 - Arithmetic Computation is designed to measure a pupil's ability to do the fundamental arithmetical operations with whole numbers, decimals and fractions.

Test 7 - Arithmetic Problem Solving and Concepts consists of items to measure the understanding of concepts of the number system and the ability to reason in numerical situations.

Test 8 - Social Studies Information consists of items to measure the knowledge outcomes in three areas: geography, history and civics.

Test 9 - Social Studies Study Skills is a test designed to evaluate the pupil's skills in such areas as reading and interpreting maps and charts.

Test 10- Science measures information, generalizations and understandings in such areas as life science, earth science, physical science, conservation and health.

By studying the bivariate tables for each of the two grade levels for each of the ten achievement tests, it was the writer's intention to examine the concept of M.A.G.E. as based on actual data.[6] Bivariate tables are variously referred to in educational literature as correlation tables, two-way tables, double-entry tables, comparison charts, or scattergrams. Essentially, all are graphic methods of showing the relationship between two sets of variables, where one axis of the diagram

---

6 The writer requested and received permission from the Chicago Board of Education to use the bivariate tables which had been prepared by him for use in a study of the testing program.

represents one variable and the other axis represents another variable. In this particular instance the vertical axis was used for the total mental age from the C.T.T.M. Mental ages were grouped in fifteen half-year intervals from 8 - 11 and below to 15 - 0 up for 6A pupils, and from 10 - 5 and below to 17 - 0 and above for 8B pupils. The horizontal axis was used for the particular achievement test under study. For purposes of comparison with the M.A.G.E., the data from these bivariate tables the twenty-fifth, the fiftieth, and the seventy-fifth percentile grade equivalents for each of fifteen six-month age intervals for each of the tests was calculated. As a result of these calculations it was possible to prepare tables for each test indicating the twenty-fifth, fiftieth and seventy-fifth percentiles for each of the fifteen mental age intervals.

In an effort to graphically demonstrate the marked differences between the theoretical concept as proposed by the M.A.G.E. and the actual quartiles as calculated from the data, tables were prepared for the sixth grade and the eighth grade, ten for each grade level. Each table compares the actual twenty-fifth, fiftieth and seventy-fifth percentile grade equivalent with the mental age grade expectancy for the particular test and grade level. For example, in Table III, for the sub group of pupils of mental ages 13 - 0 to 13 - 5, the distribu-

tion of grade scores, or grade equivalents, has a twenty-fifth percentile point of 6.4, a fiftieth percentile point or median of 7.5, and a seventy-fifth percentile point of 9.5 as is indicated by the points plotted above $\frac{13\text{-}0}{13\text{-}5}$ and through which the ——— line, the ···· line and the ---- line are drawn. The M.A.G.E. has been included in the form of a red bar. The lower portion of the bar corresponds to the Grade Expectancy for the lower limits of the mental age interval and the upper portion corresponds to the Grade Expectancy of the upper limits of the interval.

An examination of the Word Knowledge Test for sixth grade, which is the first table, reveals several interesting facts. By following the fiftieth percentile for each of the mental age intervals it is quite obvious that achievement does increase in relation to mental ability but it does not increase at an even rate as is suggested by M.A.G.E. Further analysis of this same table would also indicate an interesting increase in the fiftieth percentile grade equivalent as you go from the interval of 13-0 - 13-5 to the interval of 13-6 - 13-11. Such a difference cannot be readily accounted for but a plausible explanation would seem to be that this unusual increase might be accounted for by the norm group of the Word Knowledge Test. An examination of this table as well as the subsequent

ones will aid the reader in seeing that the median grade equivalents do vary from test to test, and from interval to interval. Furthermore, these variations are not consistent nor do they seem to follow a particular pattern.

TABLE III

SIXTH GRADE WORD KNOWLEDGE
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

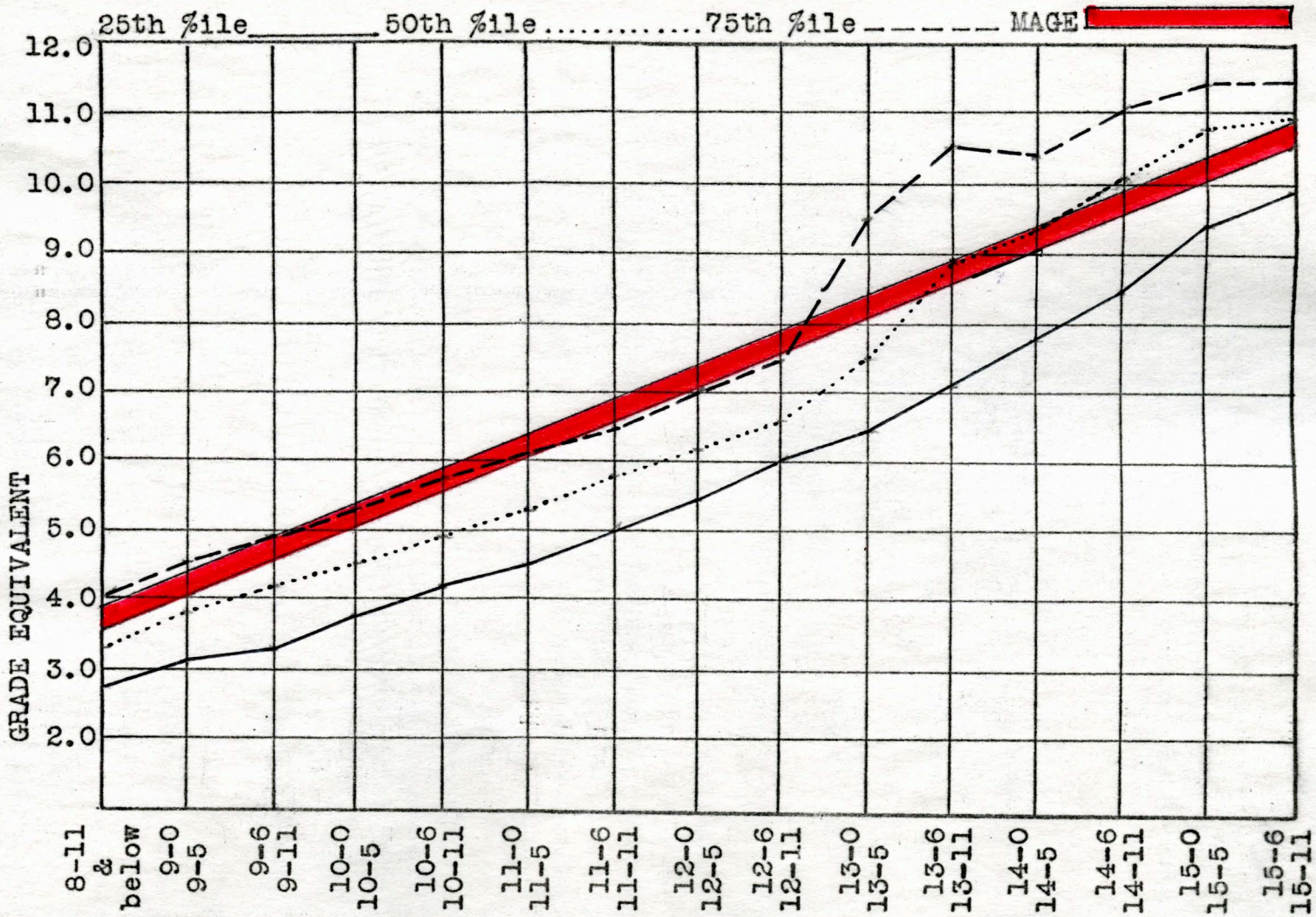


TABLE IV

SIXTH GRADE READING
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

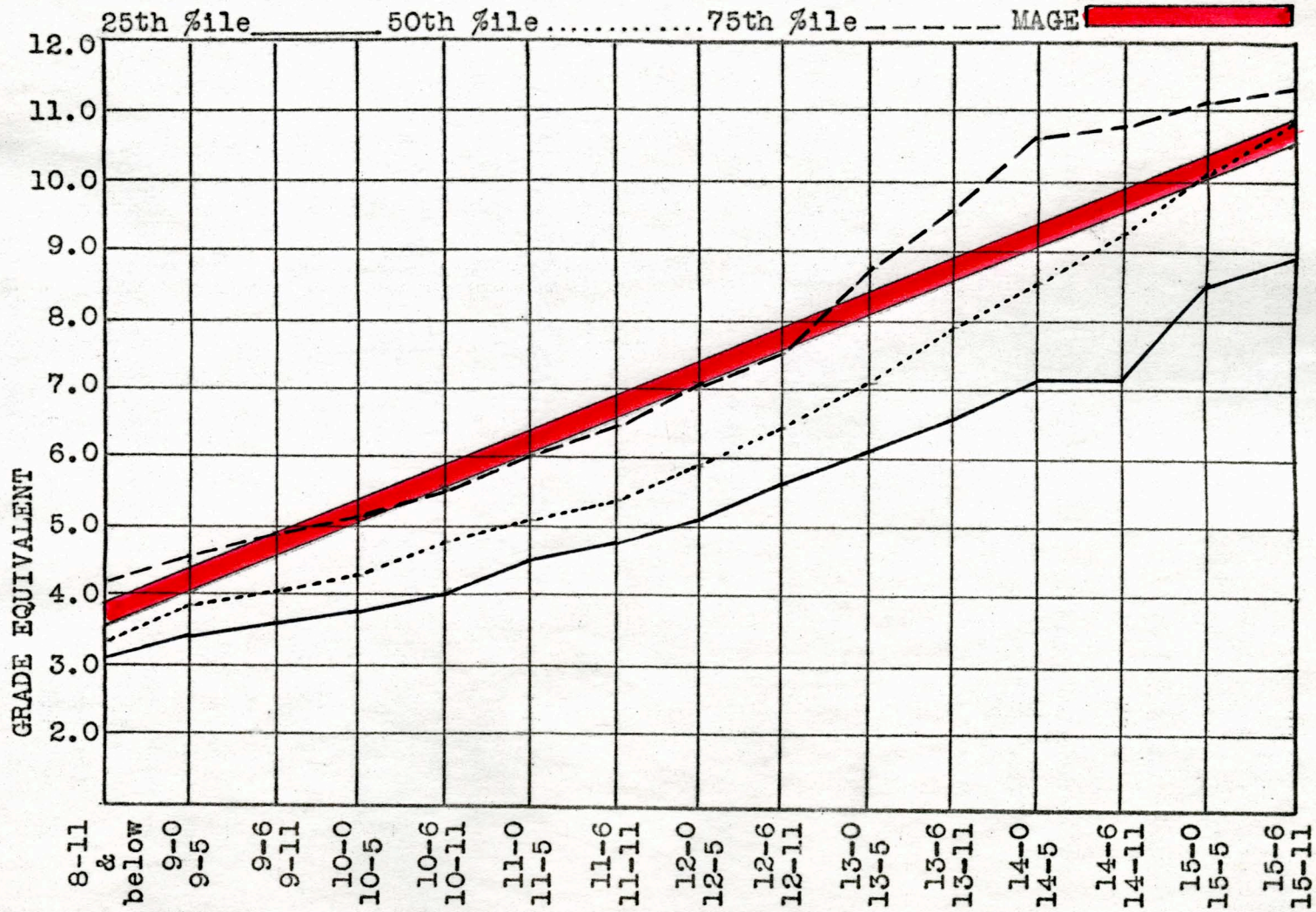


TABLE V

SIXTH GRADE SPELLING
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

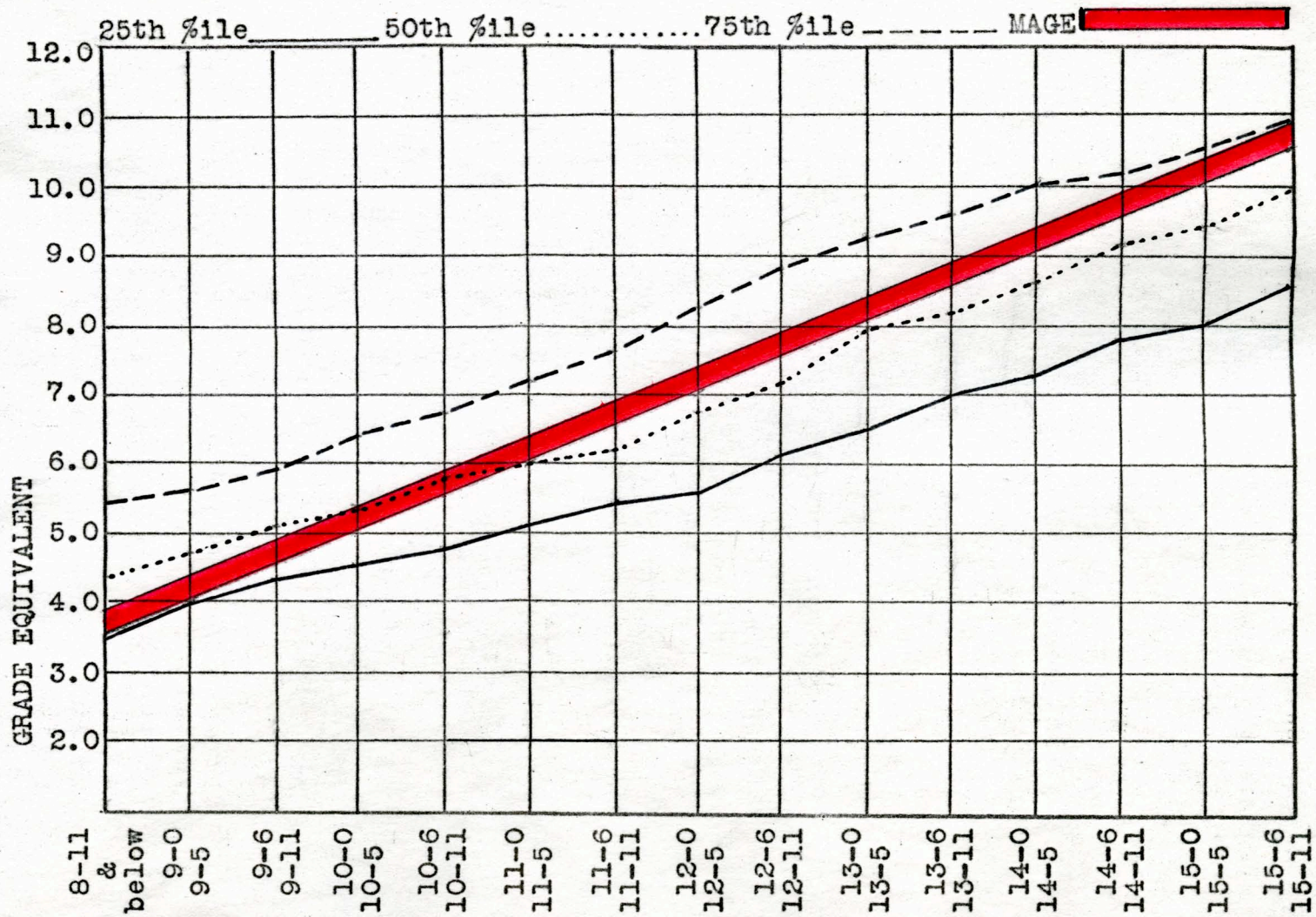


TABLE VI

SIXTH GRADE LANGUAGE
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

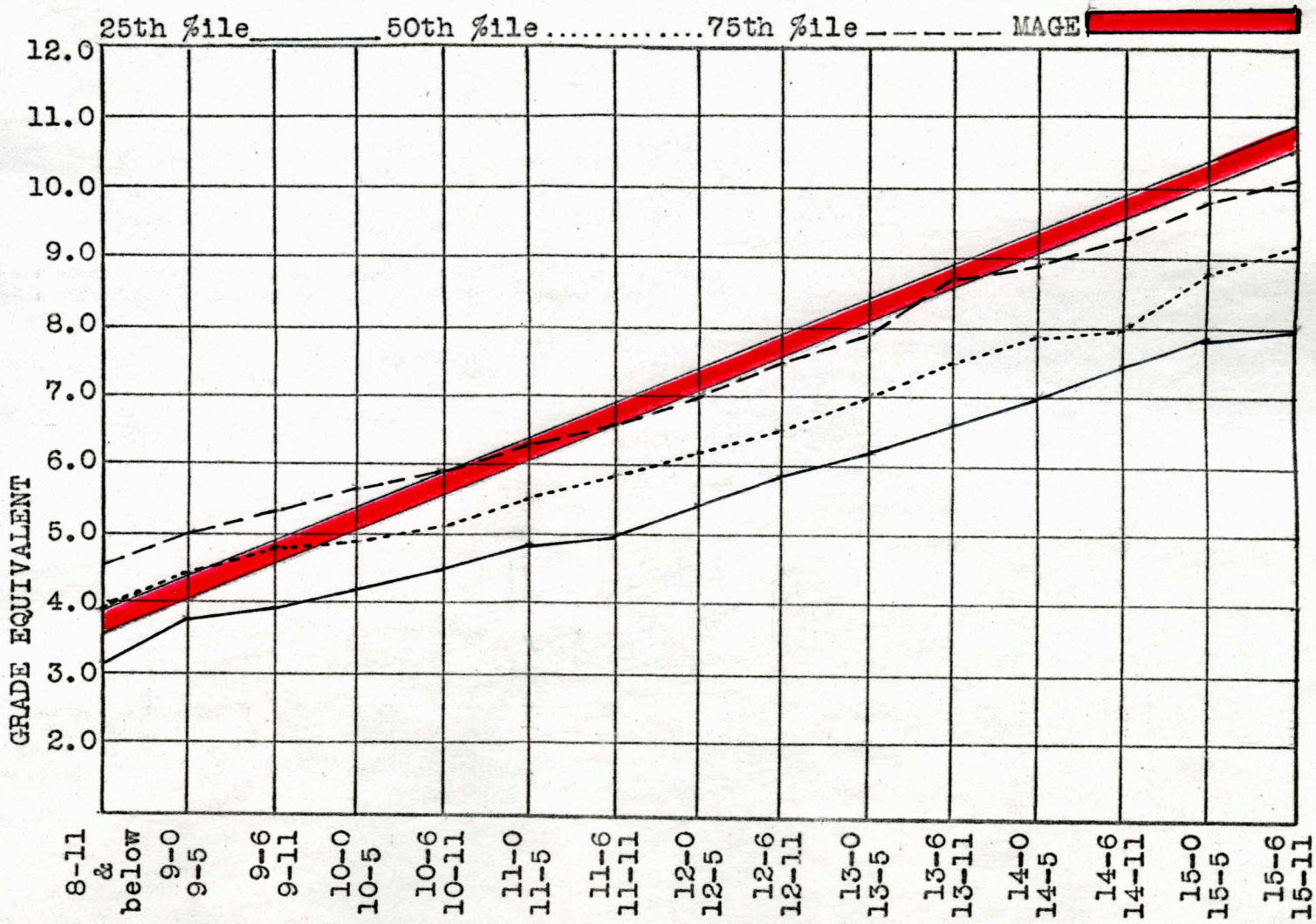


TABLE VII

SIXTH GRADE LANGUAGE STUDY SKILLS
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

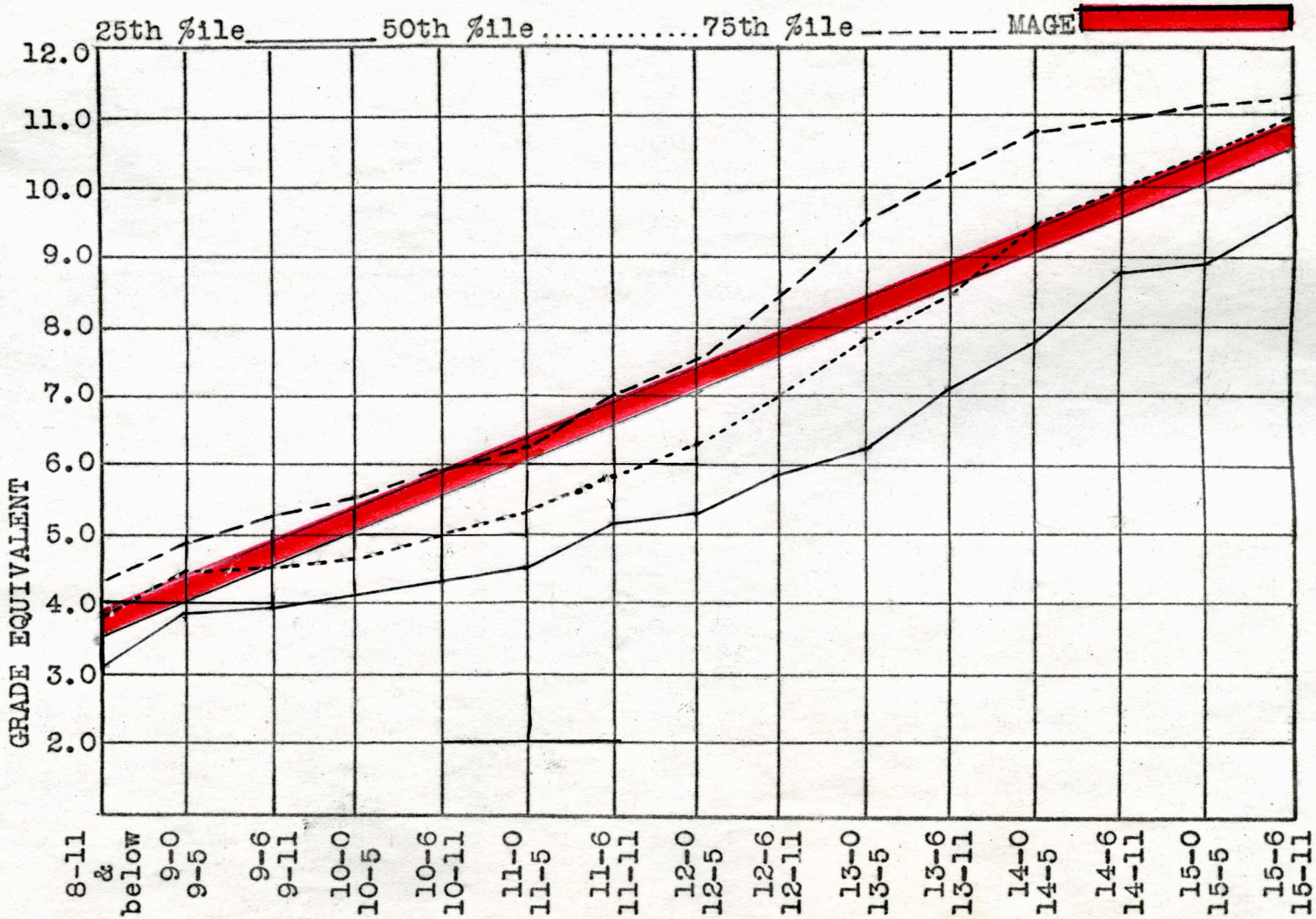


TABLE VIII

SIXTH GRADE ARITHMETIC COMPUTATION
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

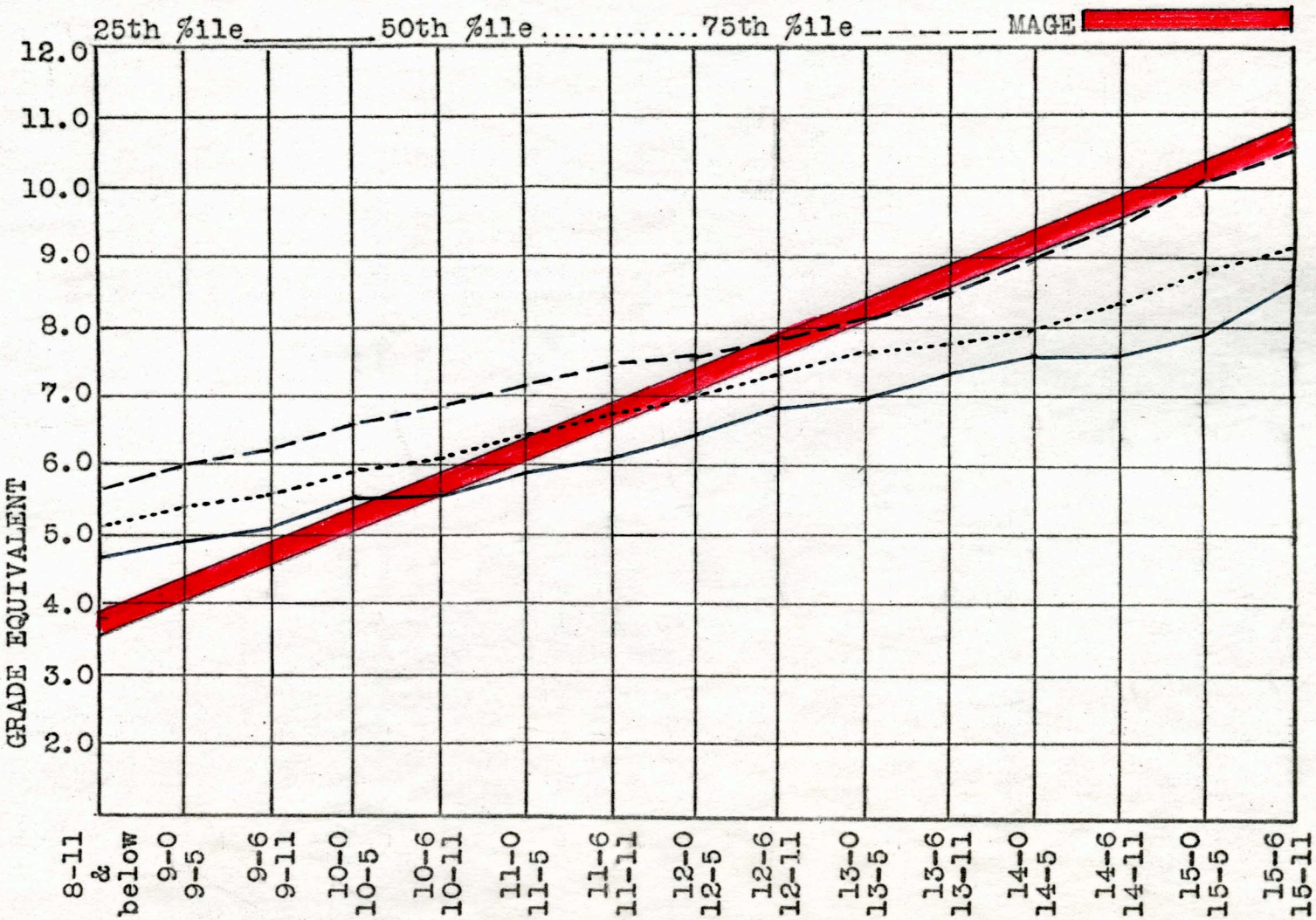


TABLE IX

SIXTH GRADE ARITHMETIC PROBLEM SOLVING
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

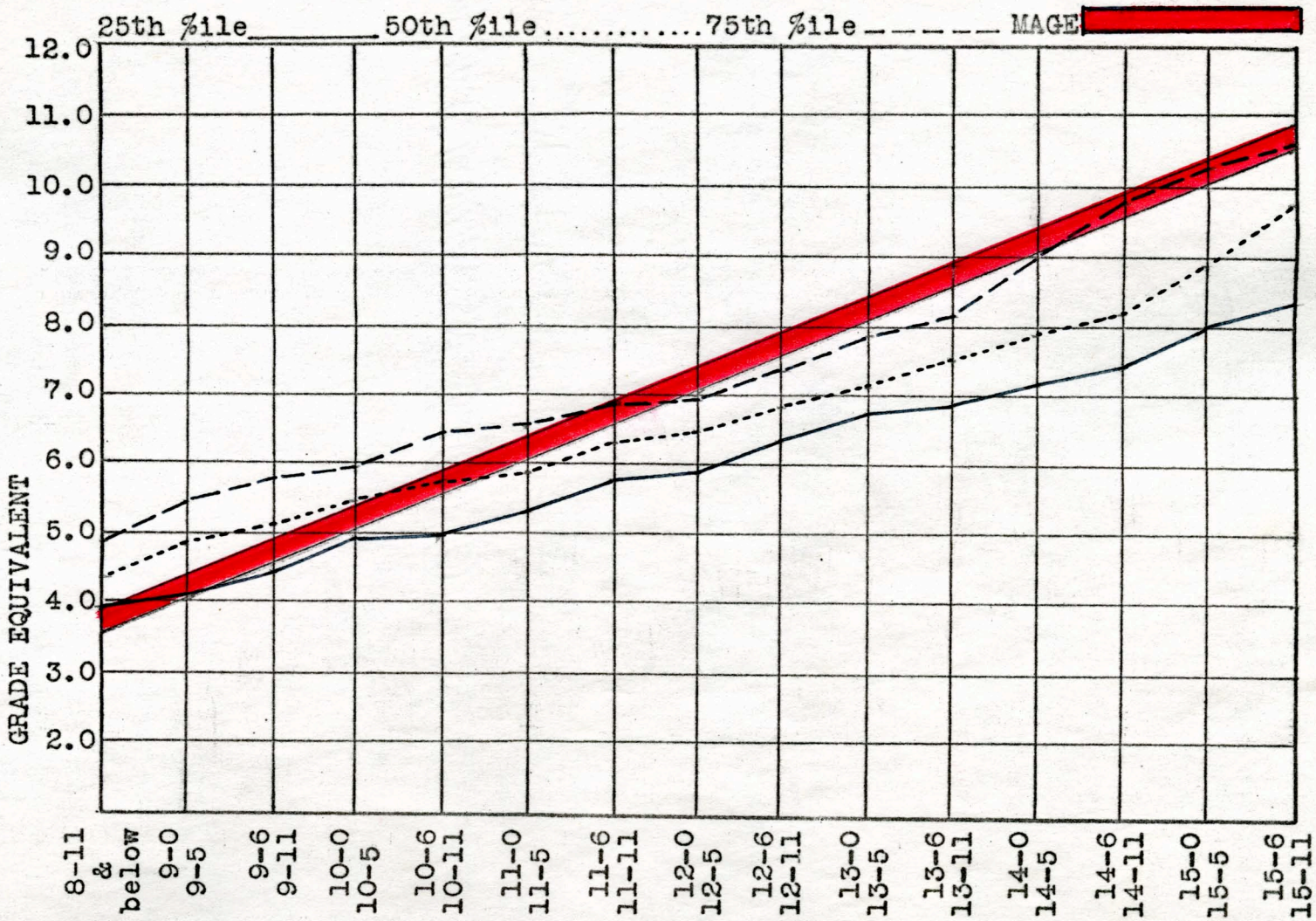


TABLE X

SIXTH GRADE SOCIAL STUDIES INFORMATION
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

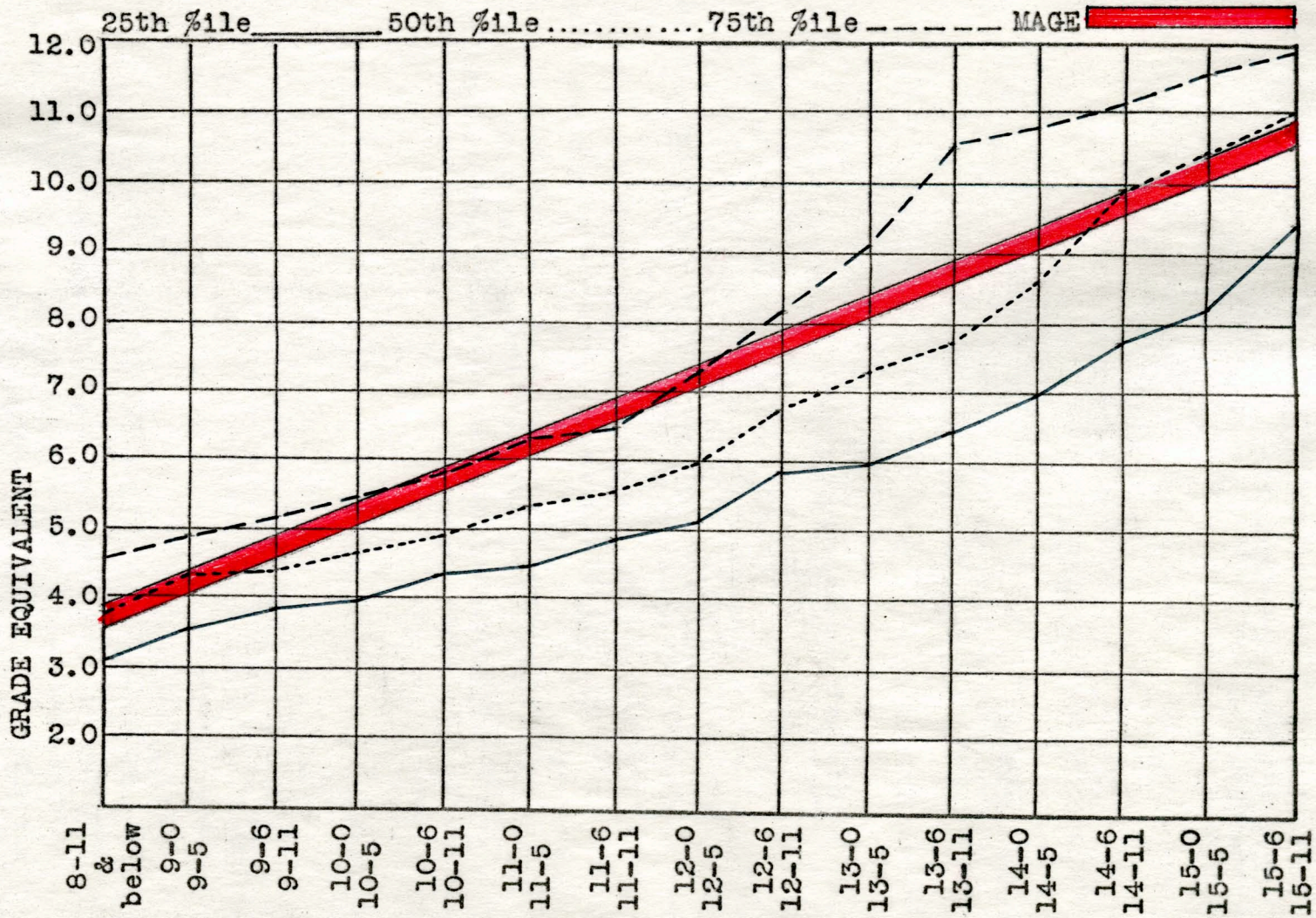


TABLE XI

SIXTH GRADE SOCIAL STUDIES STUDY SKILLS
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

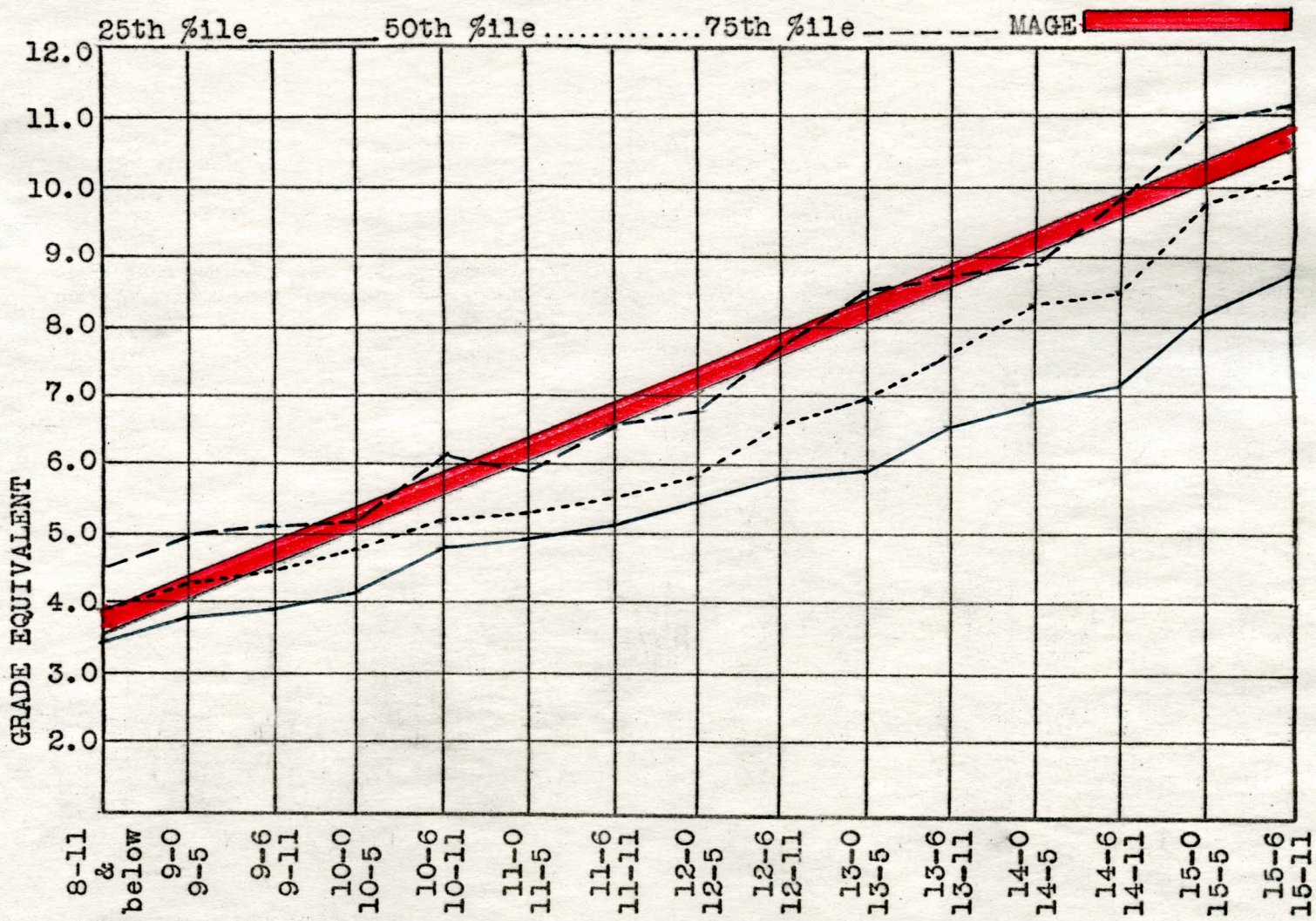


TABLE XII

SIXTH GRADE SCIENCE
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

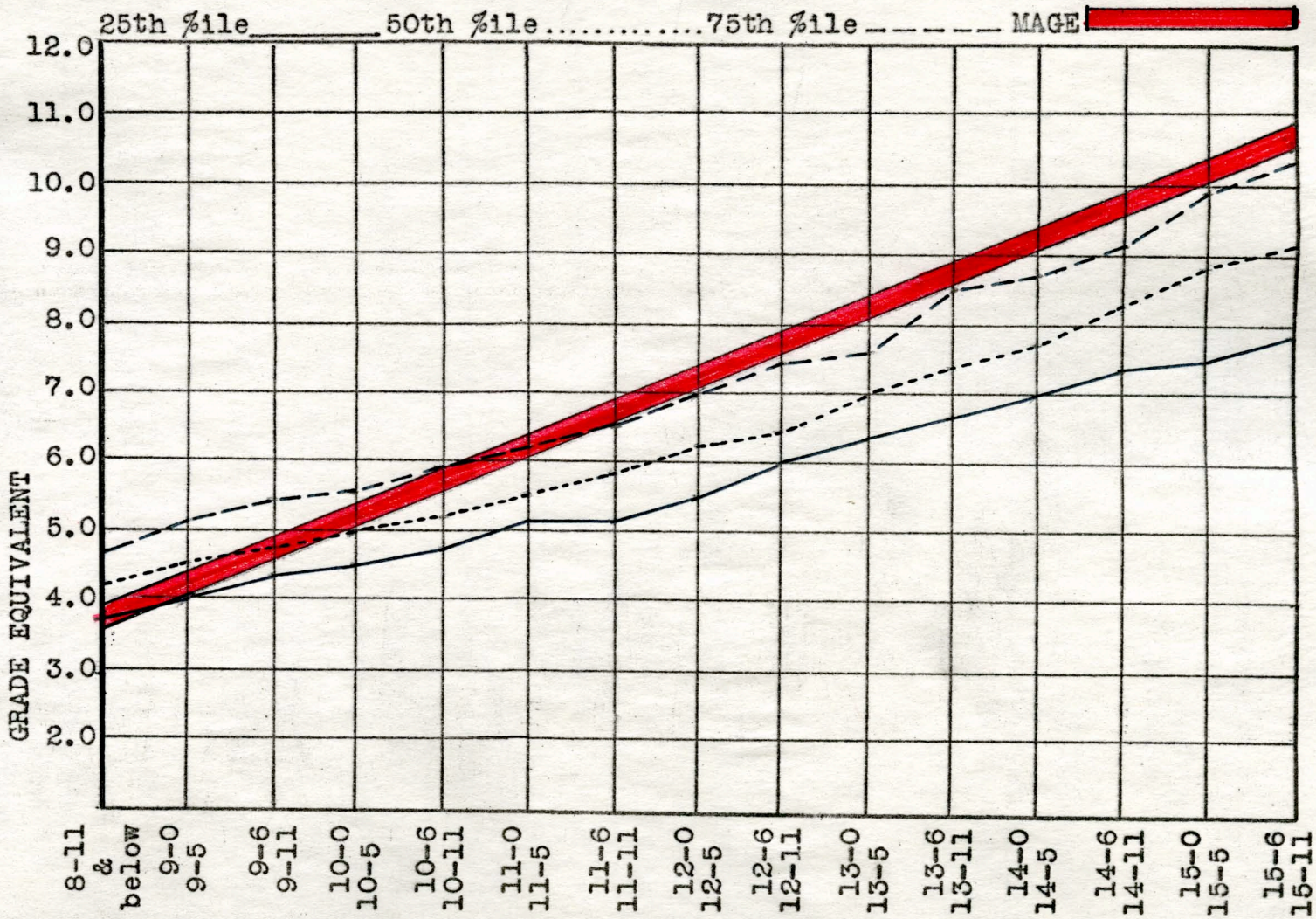


TABLE XIII

EIGHTH GRADE WORD KNOWLEDGE
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

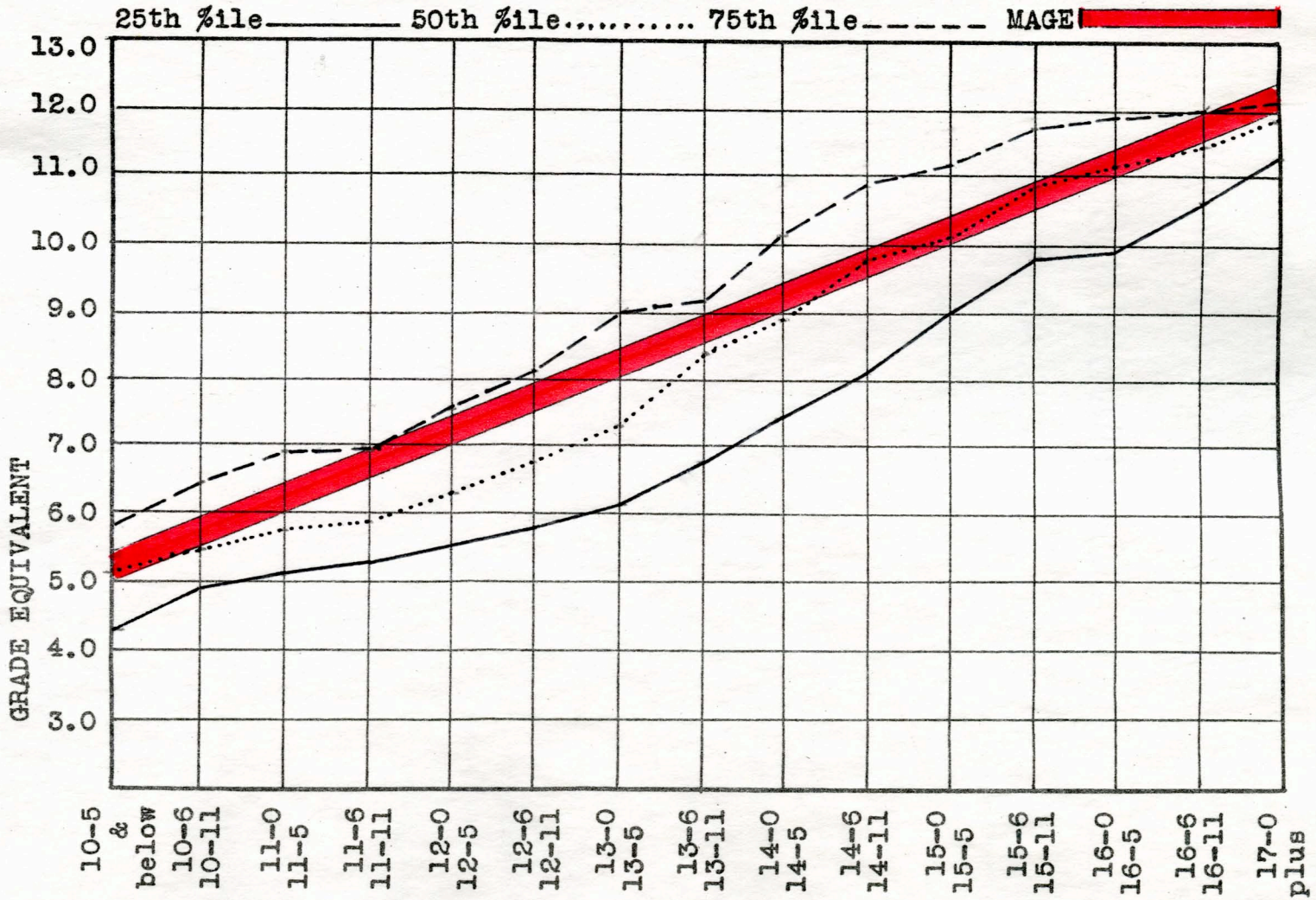


TABLE XIV

EIGHTH GRADE READING
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

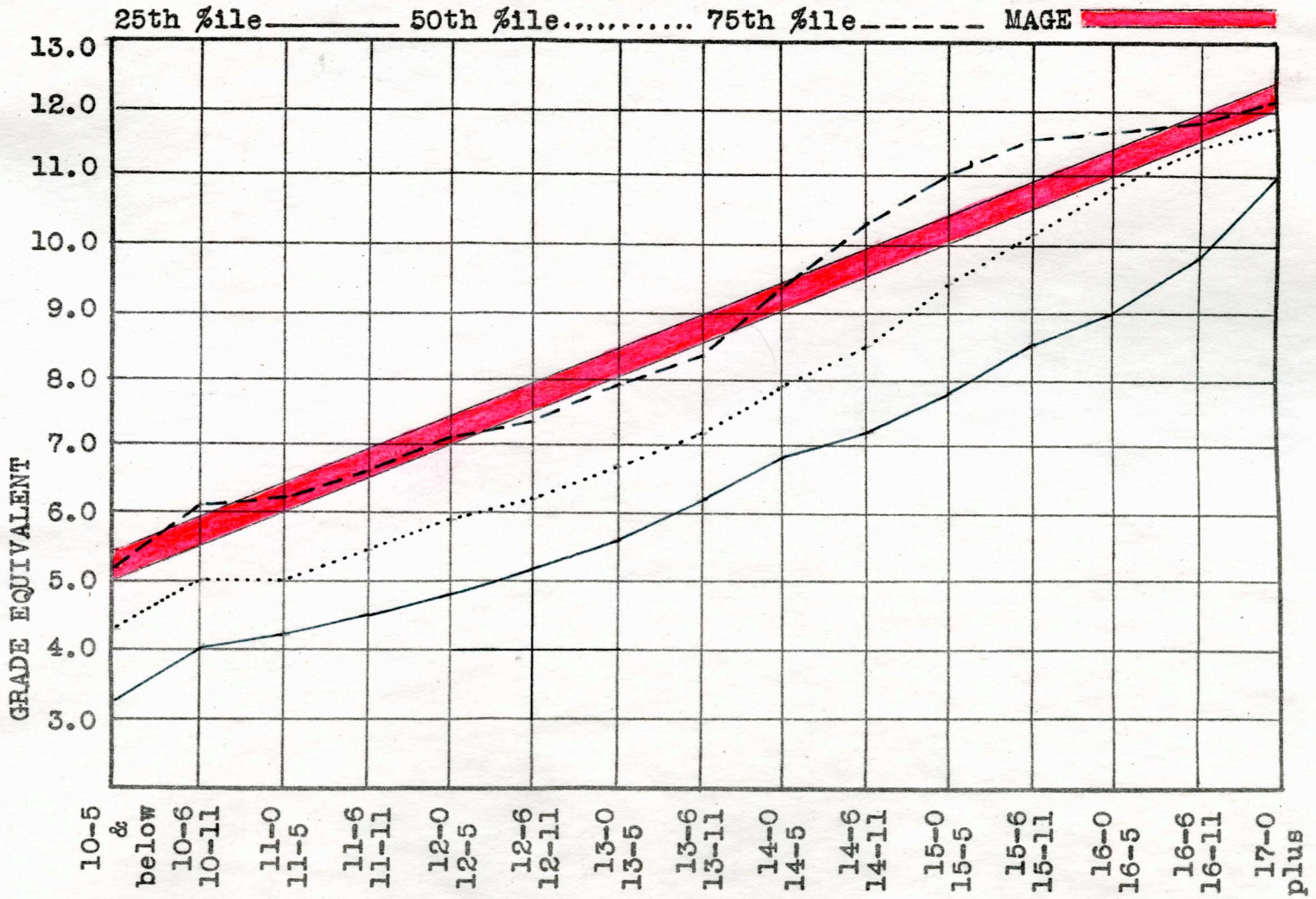


TABLE XV

EIGHTH GRADE SPELLING
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

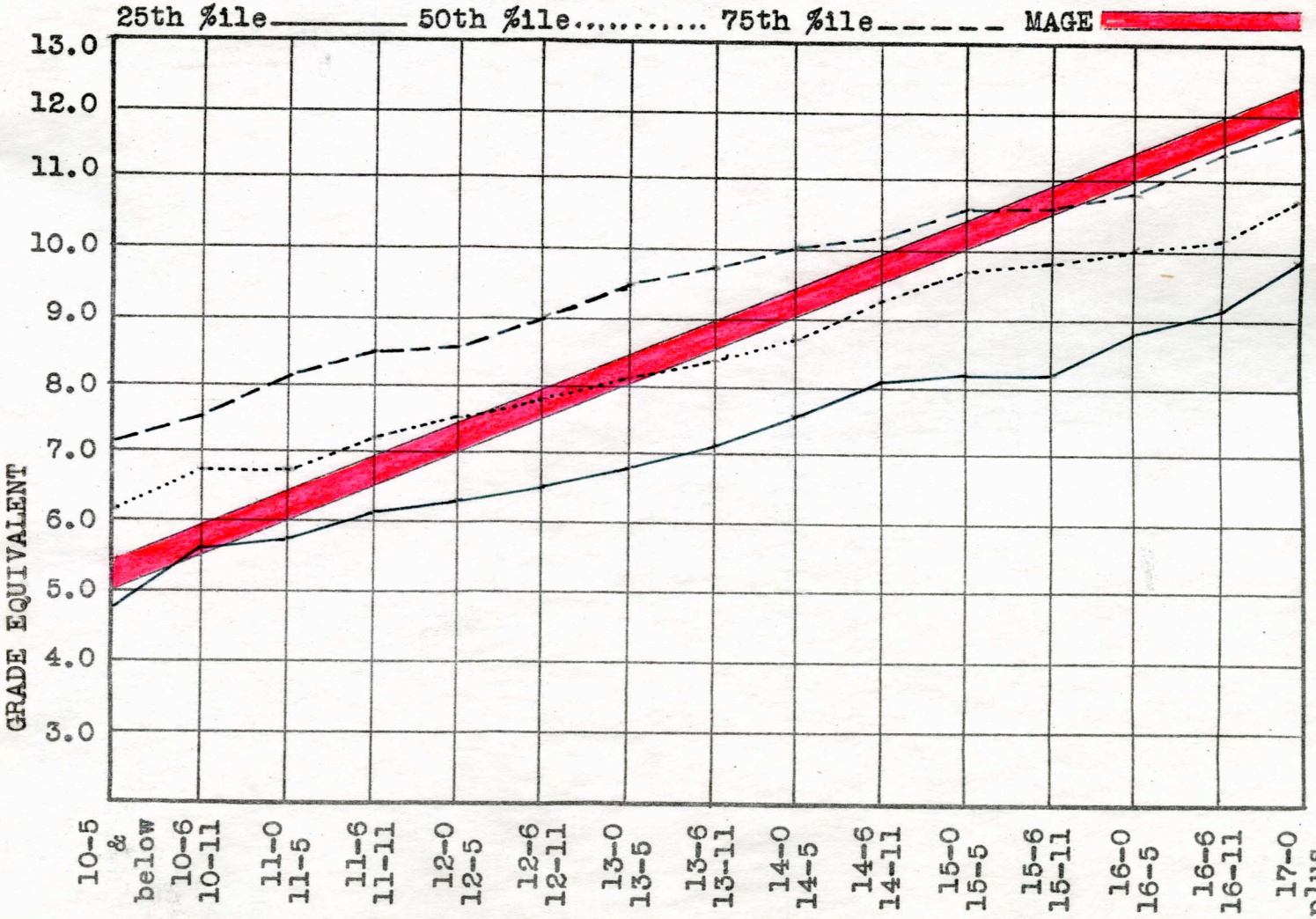


TABLE XVI

EIGHTH GRADE LANGUAGE
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

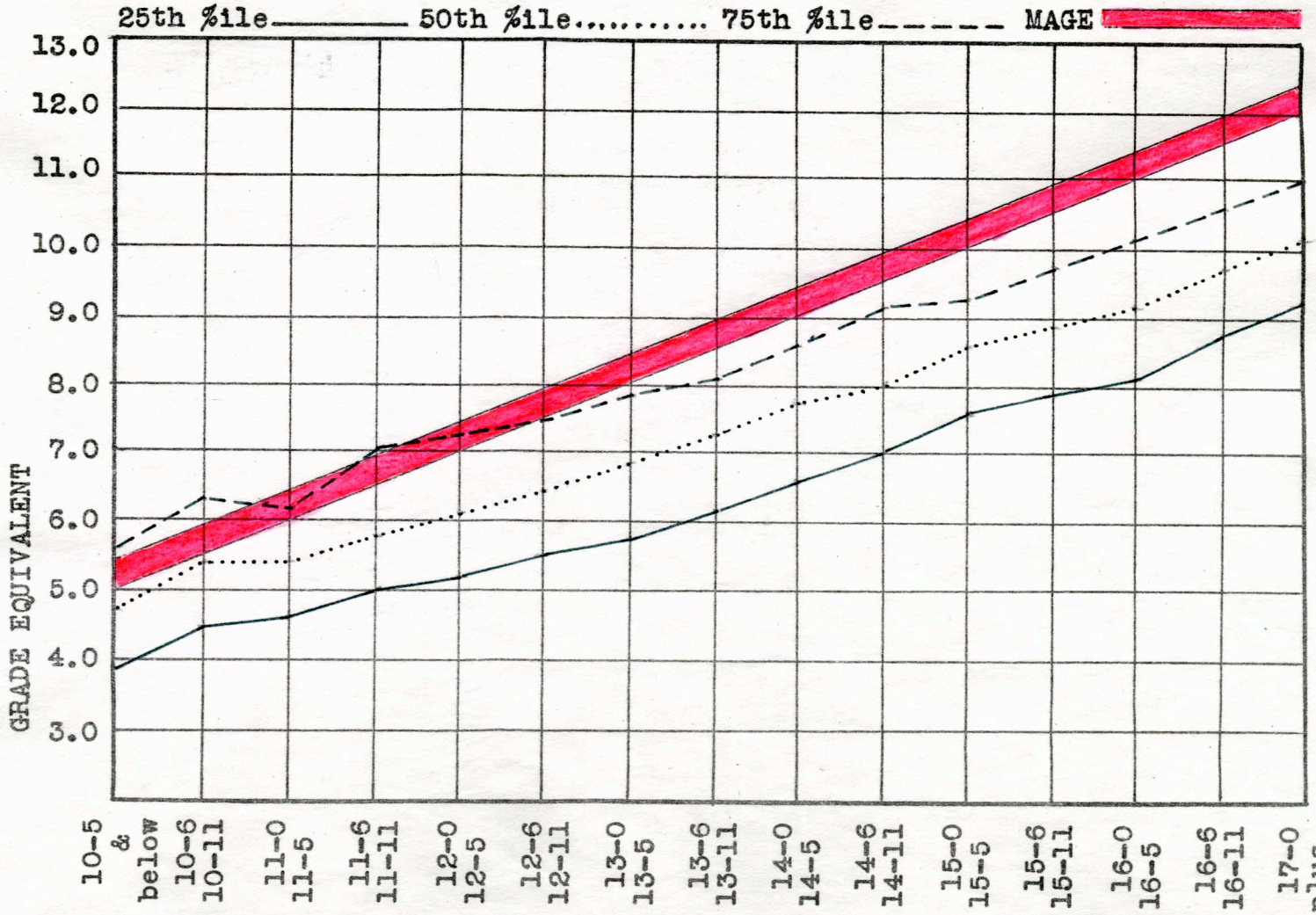


TABLE XVII

EIGHTH GRADE LANGUAGE STUDY SKILLS
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

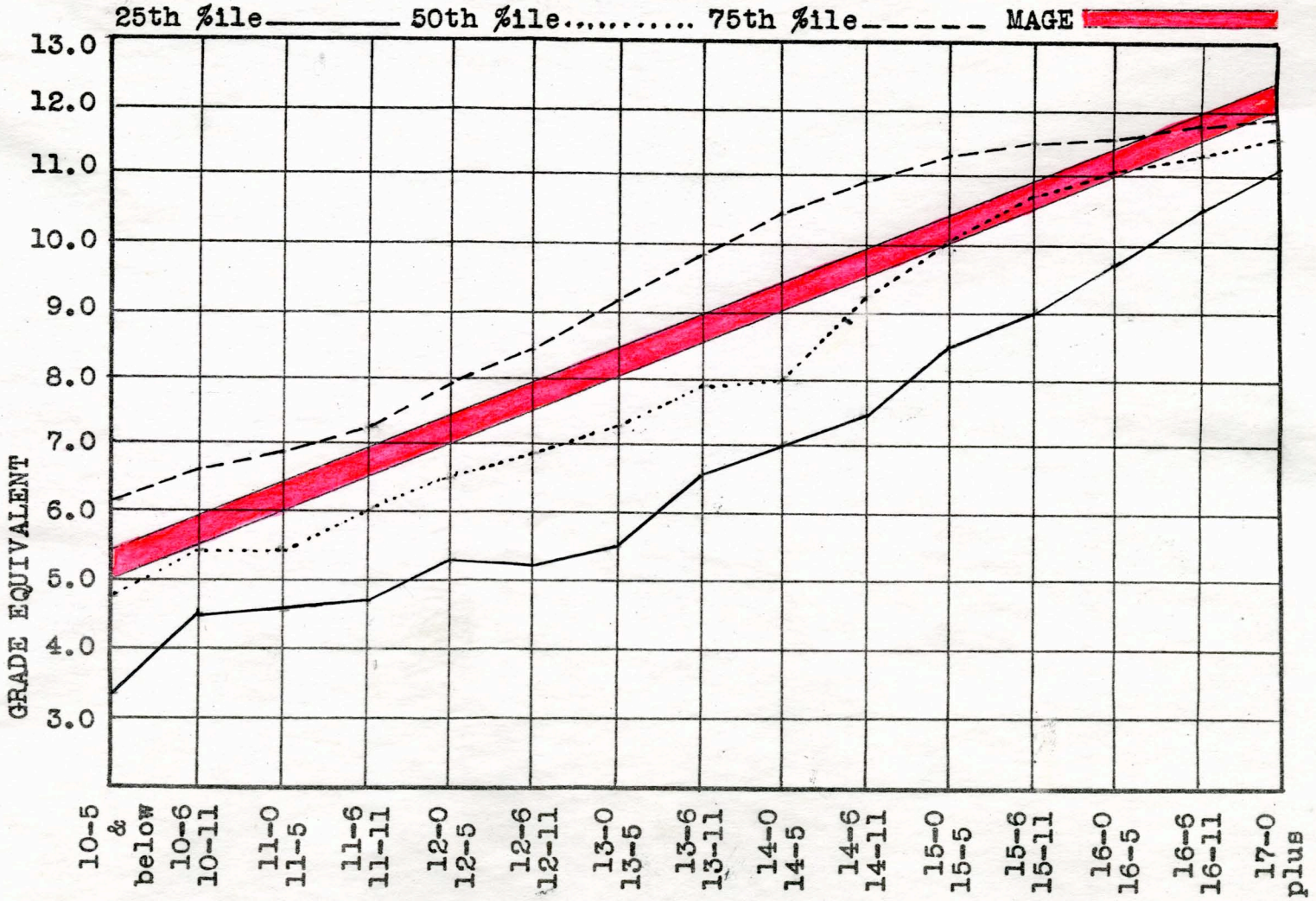


TABLE XVIII

EIGHTH GRADE ARITHMETIC COMPUTATION
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

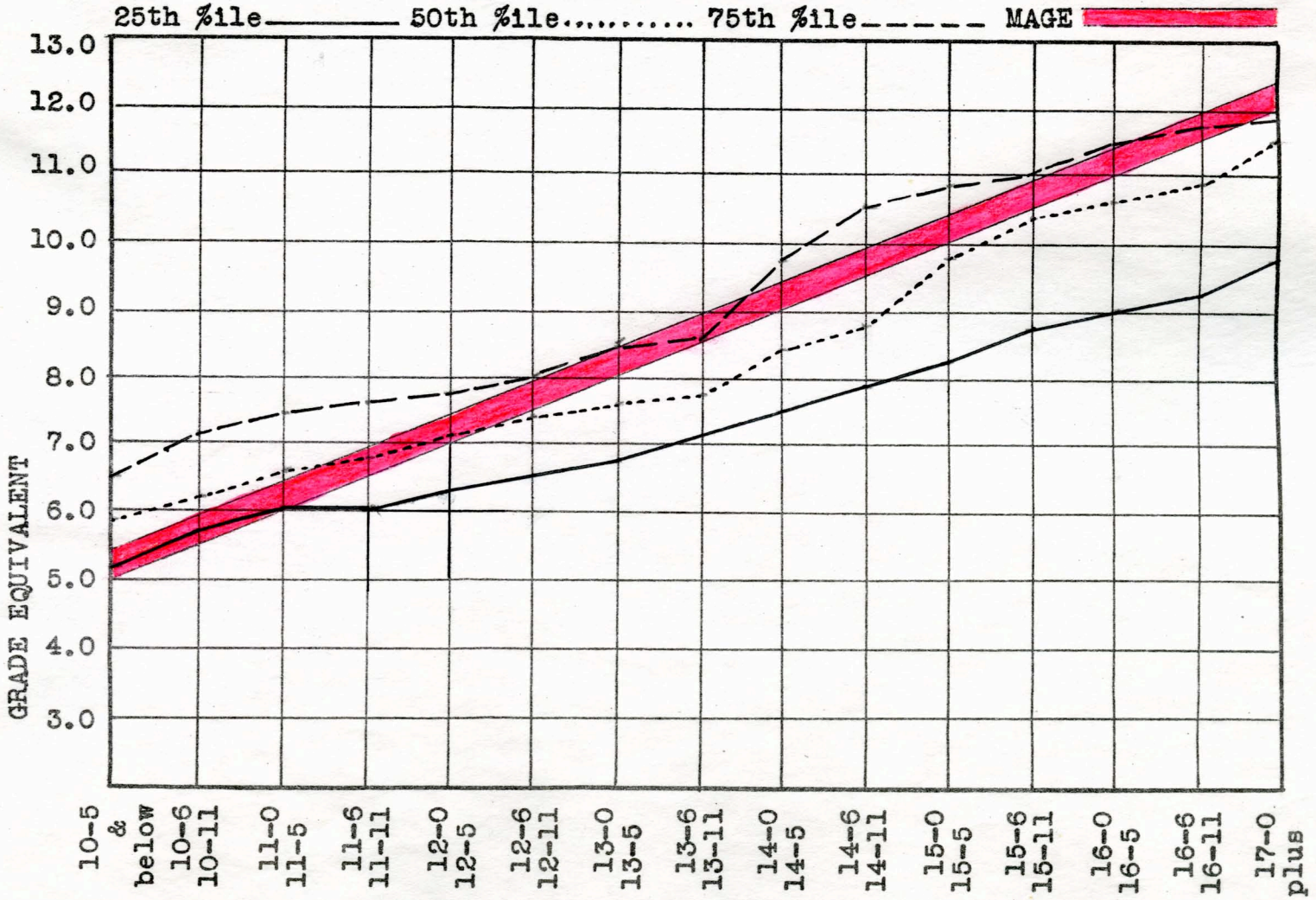


TABLE XIX

EIGHTH GRADE ARITHMETIC PROBLEM SOLVING
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

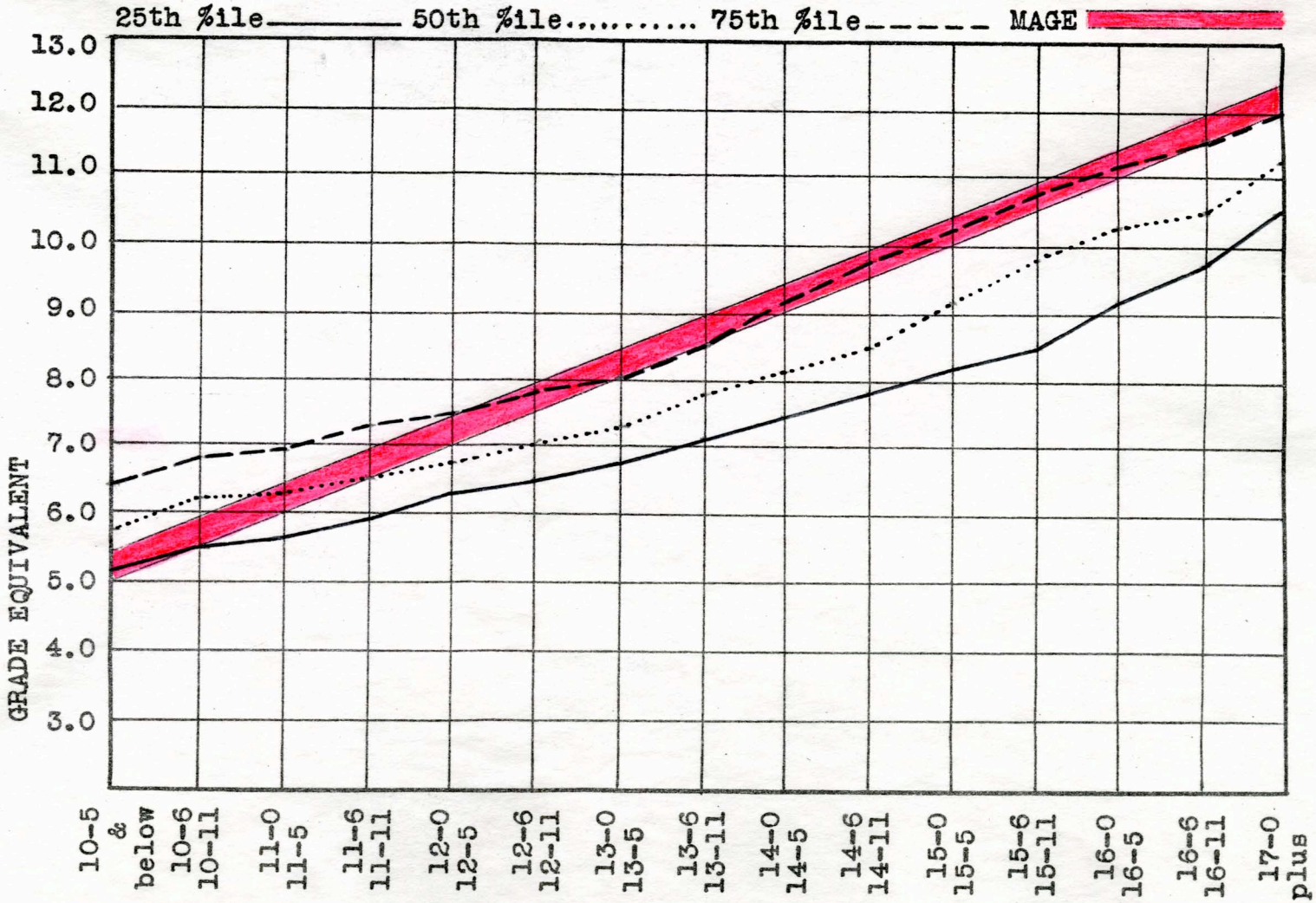


TABLE XX

EIGHTH GRADE SOCIAL STUDIES INFORMATION
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

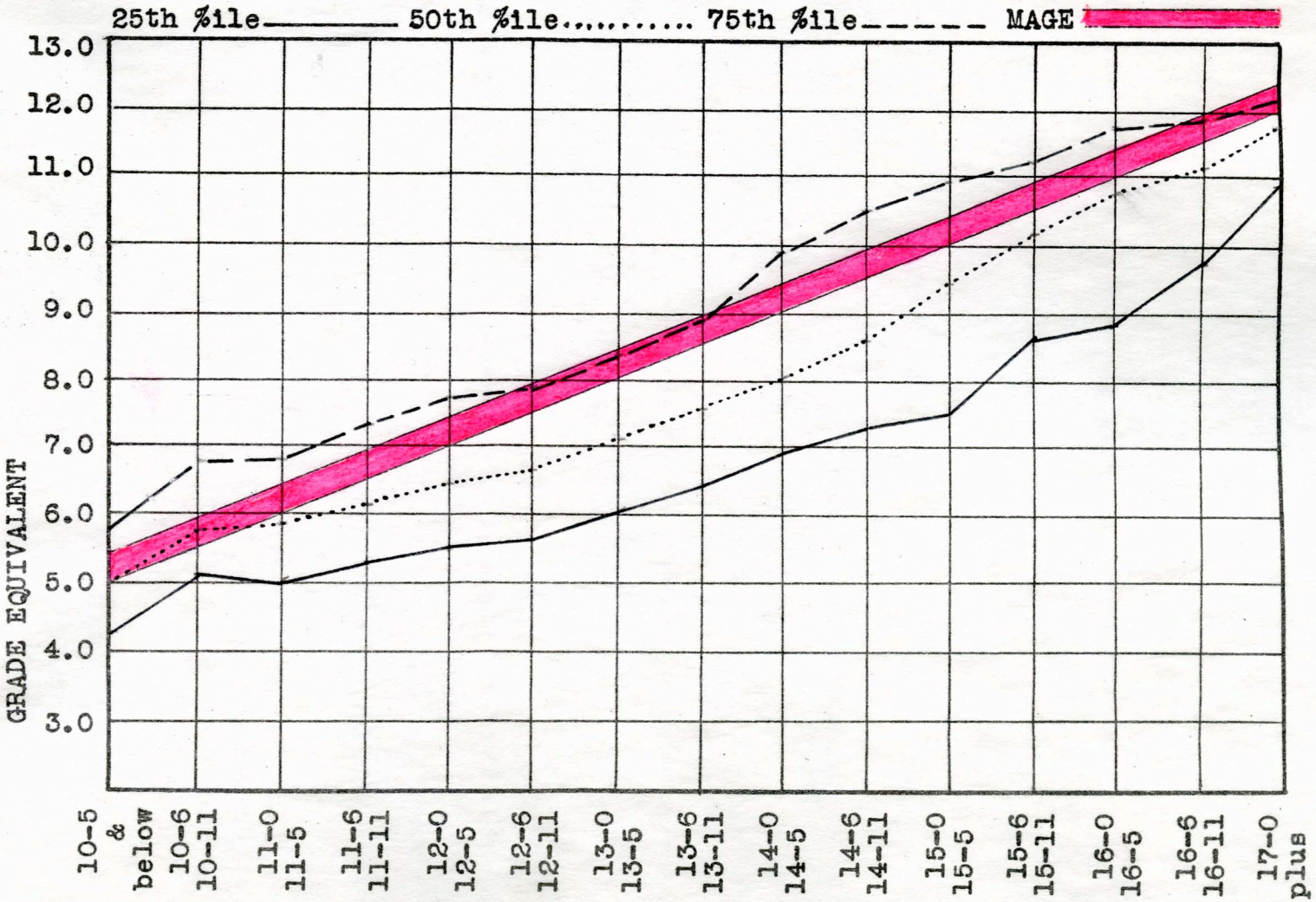


TABLE XXI

EIGHTH GRADE SOCIAL STUDIES STUDY SKILLS
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY

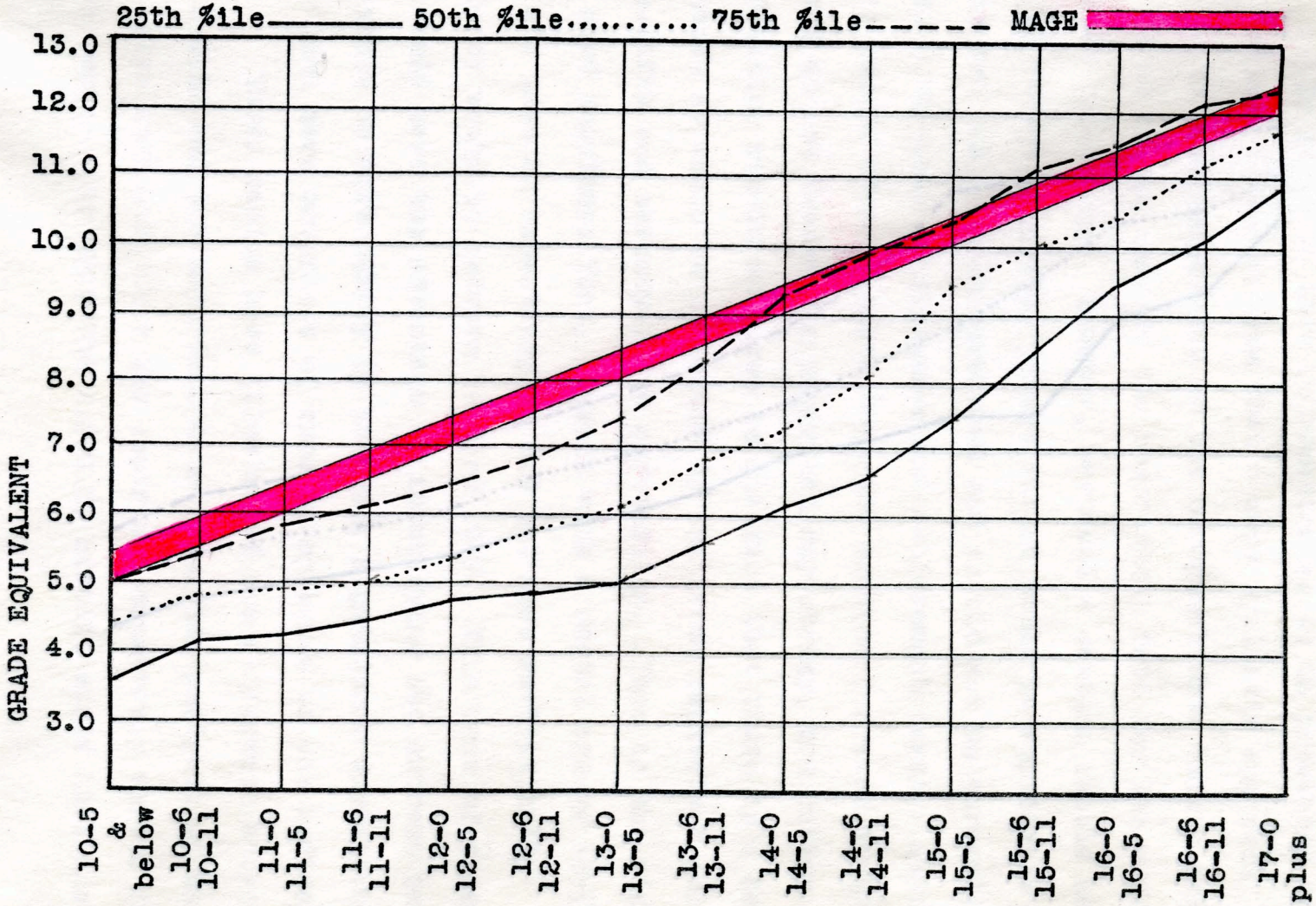
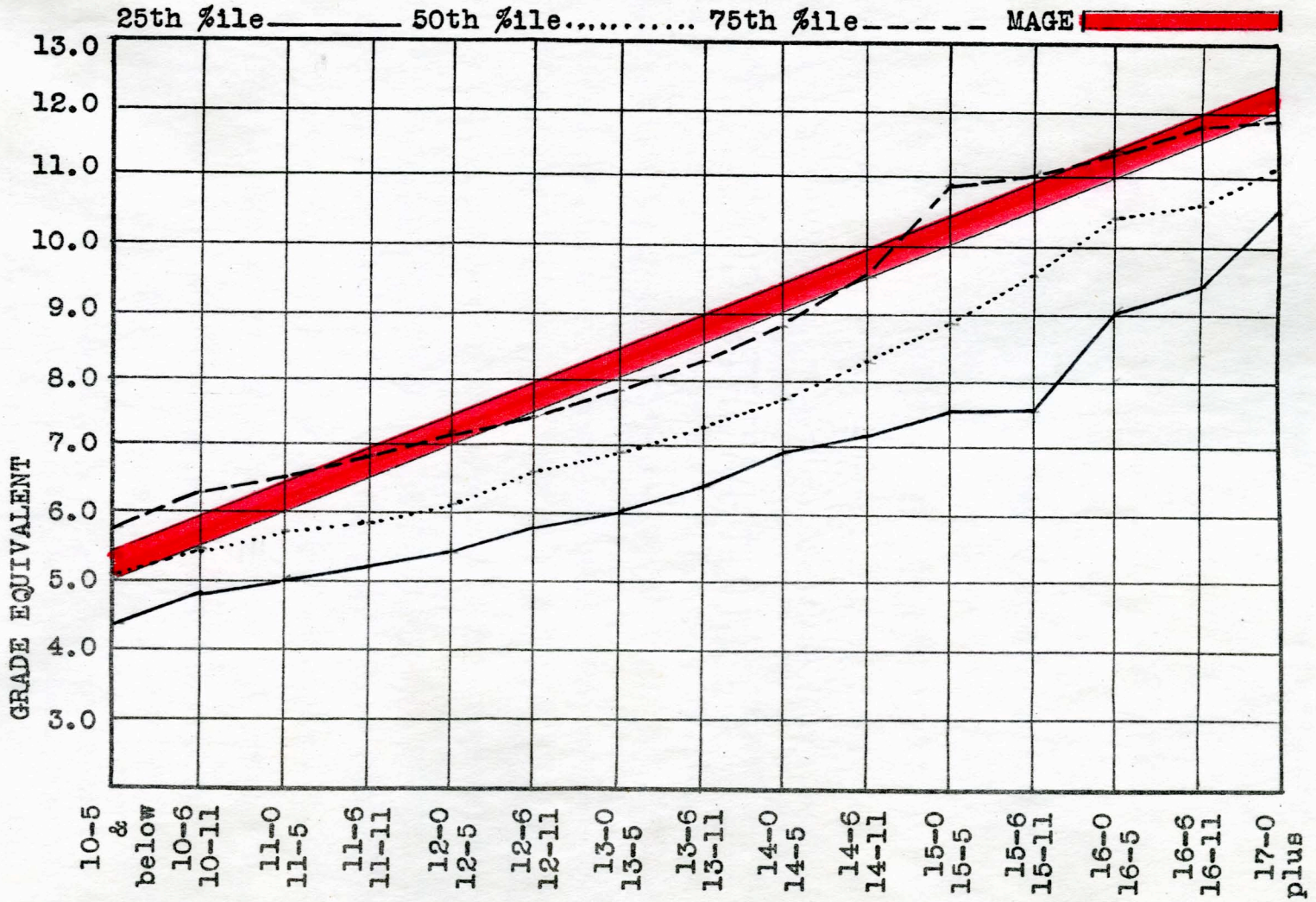


TABLE XXII

EIGHTH GRADE SCIENCE
25th, 50th, 75th PERCENTILE GRADE EQUIVALENTS
FOR 15 MENTAL AGE INTERVALS CONTRASTED WITH
MENTAL AGE GRADE EXPECTANCY



In light of this information as it is presented in these tables, it would seem that a more realistic approach would be to base expectancy on the empirical data herein presented in these tables than on the theoretical data for the mental age grade expectancy. However, particular caution should be taken in several regards. These tables would be applicable only for the two tests which were used and should not be generalized for use with other intelligence and achievement tests. Additional care should also be taken that the medians for each mental age interval do not become the empirical expectancy, but rather some consideration should be given to the fact that there is a band of expectancy as would be indicated by the grade equivalent for the twenty-fifth and seventy-fifth percentiles. Thus on the sixth grade Word Knowledge Test, a pupil with a mental age of 13-7 would not have an expectancy of 8.9 but more realistically expectancy would be between 7-1 and 10-5. This type of expectancy would have the advantage of identifying the extreme cases which would be more likely to be truly significant deviations rather than mere artifacts of expectancy based on the median.

These tables have illustrated the difference between the M.A.G.E. and the empirical data. Further analysis of these charts also indicates that the empirical expectancy varies from test to test unlike what the M.A.G.E. concept would lead one to

believe. In an effort to further examine this variation in expectancy as well as an attempt to study the possible "overlap" between the ten achievement tests and the intelligence test, the writer undertook a simple type of factor analysis.

Factor Analysis of 6A - 8B Tests

Factor analysis is the result of the psychologists' effort to reduce the great number of possible variables from different tests and other methods of measurement into a more intelligible framework. It is a way of getting an overview of a large number of correlation coefficients to see if the common variance which they express, as measured in pairs of variables, can be described in broader terms. Factor analysis can also be used to reduce the confusion which test results produce when the same ability is given different names in different tests.

At the present time there are a variety of methods which have been proposed to be followed in doing a factor analysis.[7] In general the British have accepted the leadership of Cyril Burt, whereas in this country most studies have followed the direction offered by L. L. Thurstone. Guilford points out that if the purpose of the factor analysis is to reduce the

---

7 Dael Wolfe, "Factor Analysis to 1940," Psychometric Monograph No. 3, 1946.

number of variables with which one operates, as is the purpose here, almost any method of factor analysis will do.[8] As a consequence the writer selected Thurstone's centroid method of factor analysis.[9]

The Data

Table XXIII and Table XXIV present the two correlation matrices upon which this aspect of the study is based. One thousand random cases were drawn from each of the two grade levels which participated in the testing programs and the correlations were calculated by a computer.[10] It should be noted that these matrices contain eleven variables; ten from the Metropolitan Battery and one from the C.T.M.M. Only the total score for the C.T.M.M. was used instead of it as well as the Language and Non-Language. This was done since the Total score is an average of the Language and Non-Language and would represent both of them. Furthermore, the Total score is more frequently used than the other two scores for prediction purposes.

---

8 J. P. Guilford, Psychometric Methods, New York, 1954, 522.

9 L. L. Thurstone, Multiple Factor Analysis, Chicago, 1947.

10 The correlations were done by the Psychological Corporation of New York.

TABLE XXIII

INTERCORRELATIONS - METROPOLITAN BATTERY
CALIFORNIA TEST OF MENTAL MATURITY
N = 1000 GRADE 6

Test	1	2	3	4	5	6	7	8	9	10	11
2	.83										
3	.75	.66									
4	.74	.74	.67								
5	.75	.77	.65	.76							
6	.64	.65	.60	.70	.66						
7	.70	.72	.63	.73	.71	.83					
8	.75	.76	.60	.67	.72	.58	.66				
9	.68	.71	.52	.67	.70	.61	.70	.73			
10	.84	.82	.63	.70	.75	.60	.69	.80	.75		
11	.72	.71	.60	.70	.68	.66	.74	.65	.69	.70	

Code for Tests

Metropolitan Battery
1. Word Knowledge
2. Reading
3. Spelling
4. Language
5. Language Study Skills
6. Arithmetic Computation
7. Arithmetic Problem Solving
8. Social Studies Information
9. Social Studies Study Skills
10. Science

California Test Mental Maturity
11. Total

TABLE XXIV

INTERCORRELATIONS - METROPOLITAN BATTERY
CALIFORNIA TEST OF MENTAL MATURITY
N = 1000 GRADE 8

Test	1	2	3	4	5	6	7	8	9	10	11
2	.85										
3	.65	.64									
4	.73	.75	.69								
5	.73	.74	.61	.74							
6	.65	.68	.58	.74	.69						
7	.70	.72	.55	.73	.69	.84					
8	.75	.78	.55	.66	.63	.64	.67				
9	.69	.72	.48	.70	.68	.70	.77	.67			
10	.76	.75	.46	.64	.64	.62	.71	.74	.70		
11	.71	.69	.50	.66	.61	.67	.73	.63	.70	.67	

Code for Tests

Metropolitan Battery
1. Word Knowledge
2. Reading
3. Spelling
4. Language
5. Language Study Skills
6. Arithmetic Computation
7. Arithmetic Problem Solving
8. Social Studies Information
9. Social Studies Study Skills
10. Science

California Test Mental Maturity
11. Total

The steps followed in the removal of a factor according to Thurstone's centroid method are as follows:

1. Reflect the correlation matrix into the upper half of the matrix.
2. Enter communalities in the diagonal cells.
3. Sum the columns and sum the rows. These should be equal.
4. Sum the sums of the rows and columns. Once again they should be equal.
5. Find the reciprocal of the square root of the quantity found in 4.
6. Multiply each column sum by the quantity found in 5.
7. Calculate the residuals for each test after the first factor is removed.

TABLE XXV

FACTOR LOADINGS

Test	(6th Grade) Factor Loading	(6th Grade) Rank Order	(8th Grade) Factor Loading	(8th Grade) Rank Order
1 Word Knowledge	.893	1	.886	2
2 Reading	.887	2	.889	1
3 Spelling	.756	11	.703	11
4 Language	.850	6	.854	4
5 Language Study Skills	.856	4	.818	8
6 Arithmetic Computation	.798	10	.840	5
7 Arithmetic Problem Solving	.851	5	.865	3
8 Social Studies Information	.837	7	.819	7
9 Social Studies Study Skills	.814	9	.825	6
10 Science	.865	3	.809	9
11 Total Mental-Age	.822	8	.801	10

An examination of the above table indicates this first factor (x) accounts for much of the original intercorrelations between tests. Just how much it does account for can be seen when the first-factor residuals are examined in the following two tables.

TABLE XXVI

FIRST FACTOR RESIDUALS
SIXTH GRADE

Test	1	2	3	4	5	6	7	8	9	10	11
2	.038										
3	.075	-.011									
4	-.019	-.014	.027								
5	-.014	.011	.003	.032							
6	-.073	-.058	-.003	.022	-.023						
7	-.060	-.035	-.013	.007	-.018	.151					
8	.003	.018	-.033	-.041	-.004	-.088	-.052				
9	-.047	-.012	-.095	-.022	.003	-.040	.007	.049			
10	.068	.053	-.024	-.035	.010	-.090	-.046	.076	.046		
11	-.014	-.019	.021	.001	-.024	.004	.040	-.038	.021	-.011	

Code for Tests

Metropolitan Battery
1. Word Knowledge
2. Reading
3. Spelling
4. Language
5. Language Study Skills
6. Arithmetic Computation
7. Arithmetic Problem Solving
8. Social Studies Information
9. Social Studies Study Skills
10. Science

California Test Mental Maturity
11. Total

TABLE XXVII

FIRST FACTOR RESIDUALS
EIGHTH GRADE

Test	1	2	3	4	5	6	7	8	9	10	11
2	.063										
3	.028	.015									
4	-.027	-.009	.090								
5	.005	.013	.035	.041							
6	-.094	-.066	-.010	.023	.003						
7	-.066	-.049	-.058	-.009	-.018	.113					
8	.024	.052	-.026	-.039	-.040	-.048	-.038				
9	-.041	-.013	-.100	-.004	-.005	.007	.056	-.006			
10	.043	.031	-.109	-.051	-.022	-.060	.010	.077	.033		
11	.001	-.022	-.063	-.024	-.045	-.003	.037	-.026	.039	.022	

Code for Tests

Metropolitan Battery
1. Word Knowledge
2. Reading
3. Spelling
4. Language
5. Language Study Skills
6. Arithmetic Computation
7. Arithmetic Problem Solving
8. Social Studies Information
9. Social Studies Study Skills
10. Science

California Test Mental Maturity
11. Total

The writer did not attempt to remove any more factors than the one because for all practical purposes the removal of this first one left such small residuals. Furthermore, the purpose of the factor analysis was not to identify all of the factors, but rather to determine insofar as possible the extent of the overlap between achievement tests and in particular the extent of the overlap between the intelligence test and the achievement test. The removal of the first factor has indicated that the extent of its influence on the various achievement tests to the point that only two intercorrelation among the achievement tests, for sixth and eighth grade, are above .10. These were the intercorrelations between the arithmetic computation and arithmetic problem solving which were .15 for sixth grade and .11 for eighth grade.

In order to help the reader in his interpretation of the two tables presenting the residuals, the following table is presented.

TABLE XXVIII

SUMMARY OF RESIDUALS FOR
SIXTH AND EIGHTH GRADE INTERCORRELATIONS

Residual Interval	6th Grade	8th Grade
.11 - .15	1	1
.06 - .10	3	2
.00 - .05	22	25
-.01 - .05	24	20
-.06 - .10	5	7
	55	55

The extent of the "overlap" between these achievement tests and this intelligence test raises the question of whether these tests are really measuring different abilities. If the overlap is as great as this aspect of the study has shown it is quite conceivable that the attempt to relate measures of ability to measures of achievement is not a practice to be advocated since these measures are different only in name.

CHAPTER V

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Summary

The purpose of this research has been threefold: (1) to investigate and examine some of the methods which have been utilized to relate measures of intelligence to measures of achievement as revealed by the literature, (2) to determine the current practices of relating such measures by a survey of school systems with enrollments over 200,000, and (3) to examine the practice of relating intelligence and achievement test results through the use of an age-grade expectancy table for the purposes of determining the limitations of such a practice. Each of these points will be considered in the order in which they appear.

An examination of the practices which have been utilized to relate measures of intelligence to measures of achievement has revealed that the accomplishment quotient technique and the use of scattergrams have been the two most frequently written about methods. The amount of literature written about the accomplishment quotient technique is nothing short of volu-

minous. From its inception in the 1920's to the present its advocates have been most enthusiastic in their support, in spite of the many published studies and articles conducted showing that the use of the accomplishment quotient is highly questionable.

Within the last two decades several other practices of relating measures of intelligence to measures of achievment have been suggested: the scatter diagram technique, Durost's modal age norms method and a similar method proposed by William Shanner. The advocates of the scatter diagram point out that the scatter diagram is a graphic method of showing the relationship between measures of intelligence and measures of achievement. One of the inherent weaknesses as pointed out by Howard Bowman is that even though a pupil is indicated on the scatter diagram as being a deviate, this does not necessarily mean that this could not be due to the error of the tests themselves.

W. Durost and William Shanner both recognized one of the inherent weaknesses in attempting to relate measures of intelligence to measures of achievement. They recognized that unless both types of tests had been standardized on the same population that it was possible that some of the differences would be based on the norm group rather than true differences. In an effort to resolve this problem Durost suggested modal age norms which would insure comparing students who for all practical pur-

poses had progressed through school at a normal rate. Shanner would have us calculate mean achievement scores of a large number of nation-wide samples of pupils homogeneous with respect to grade, age and mental ability. Individuals would then be compared with pupils in the sampling group most nearly the same grade, age and mental ability.

In order to gather information regarding the current practices of relating mental ability test data to achievement test data, a questionnaire was prepared and sent to the fifty-six cities with public school enrollments of two hundred thousand pupils and over. The forty-eight usable questionnaires were studied in an attempt to determine (1) the frequency with which both mental ability and achievement tests are administered in the same grade, and (2) the methods used to relate the two types of measures. This study revealed that forty-three out of a possible forty-eight cities administered mental ability and achievement tests in the same year at least once. Eighteen cities did it as frequently as three times and three cities did it as often as five times. An examination of the methods which were utilized revealed that the three most frequently utilized methods were scattergrams, stanines and expectancy tables.

As a final phase of this study the author made a detailed study of the practice of relating measures of intelligence to measures of achievement through the use of the mental age

grade expectancy concept as used in Chicago.[1] This part of the study was made by analyzing the mental ability and achievement tests which were administered to the 6A and 8B pupils during the Spring of 1961. In order to examine this practice bivariate tables were prepared for each of the two grade levels for each of the ten achievement tests. The purpose of this analysis was to examine the concept of M.A.G.E. based on actual data. An examination of these charts indicated that although as mental age increased achievement also increased, achievement gains were not equal as the M.A.G.E. would indicate. Furthermore these charts graphically indicated that within any one mental age there was a wide range of achievement which could be expected.

In an effort to determine the extent to which the mental ability test and the achievement tests "overlapped" the writer undertook a simple type of factor analysis. The extent to which this overlap did exist was clearly demonstrated. After the first common factor was removed the remaining residuals were insignificant. Only two of the remaining residuals were larger than .10. These were the intercorrelations between arithmetic computation and arithmetic problem solving tests which were .15 for the 6th grade and .11 for 8th grade.

---

1 Although this concept was widely accepted when this study was first undertaken, at the present time it does not have as many supporters and is gradually being replaced by the use of stanines.

Conclusions

The attempts which have been made to relate measures of capacity to measures of achievement have been many and varied. Much of this work has suffered under the assumption and obsession that the intelligence or mental ability or scholastic aptitude tests are absolute truth. And as a consequence it has been assumed that achievement somehow ought to correspond exactly to the level of performance on the intelligence test. Laboring under this type of thinking something was wrong if the achievement did not match the "capacity." It is imperative that the intelligence test be put in its proper perspective as just another test, just another sample of behavior.[2]

Seen in this perspective the problem of relating measures of intelligence to measures of achievement is really another phase of the general problem of prediction. When measures of intelligence are related to measures of achievement part of the problem of prediction is due to the errors of measurement both in the intelligence measures as well as in the measures of achievement. With both a fallible predictor and a fallible criterion, discrepancies between the two are inevitable, if for no other reason than the measurement errors in each.

---

2 A recent publication which fully develops this concept is: Robert L. Thorndike, The Concepts of Over- and Under-achievement, New York, 1963.

This problem is further complicated when there is as sizable a correlation between the predictor and the criterion as there is between intelligence tests and achievement tests.

Relating measures of capacity to measures of achievement is further confused by the regression effect. Briefly this is the effect whereby pupils who are very high on the first measure will not be quite so high on the second and conversely the pupils who are very low on the first measure will not be quite so low on the second. This particular aspect of the problem was at least in part responsible for some of the investigators of this problem to conclude we are short-changing the able and the below average pupil is receiving a superior education. It is only if a pupil deviates significantly from predicted achievement that he may truly be an "overachiever" or "underachiever."

It is with these reservations and limitations in mind that the following recommendations are made when measures of capacity are related to measures of achievement.

Recommendations

1. Measures of capacity and measures of achievement should be recognized for what they are: samples of behavior which are subject to the same kinds of statistical error as any other test.

2. Measures of capacity and measures of achievement should not be interpreted and looked upon as separate and distinct types of tests. There is a great deal of overlap between these types of tests as was seen in the factor analysis.
3. Caution should be utilized in using grade equivalents when comparing test results from different achievement tests.
4. Mental age grade expectancy as a method used to relate measures of intelligence to measures of achievement should not be used.
5. When measures of intelligence and achievement are compared they should have been standardized on the same national population or local norms should be developed for both.
6. When relating measures of intelligence to measures of achievement it is advantageous to convert both scores to a common type of standard score such as stanines.
7. The use of stanines in relating measures of intelligence to measures of achievement has an advantage over the use of other scores since stanines are band scores.
8. If stanines are used to relate measures of

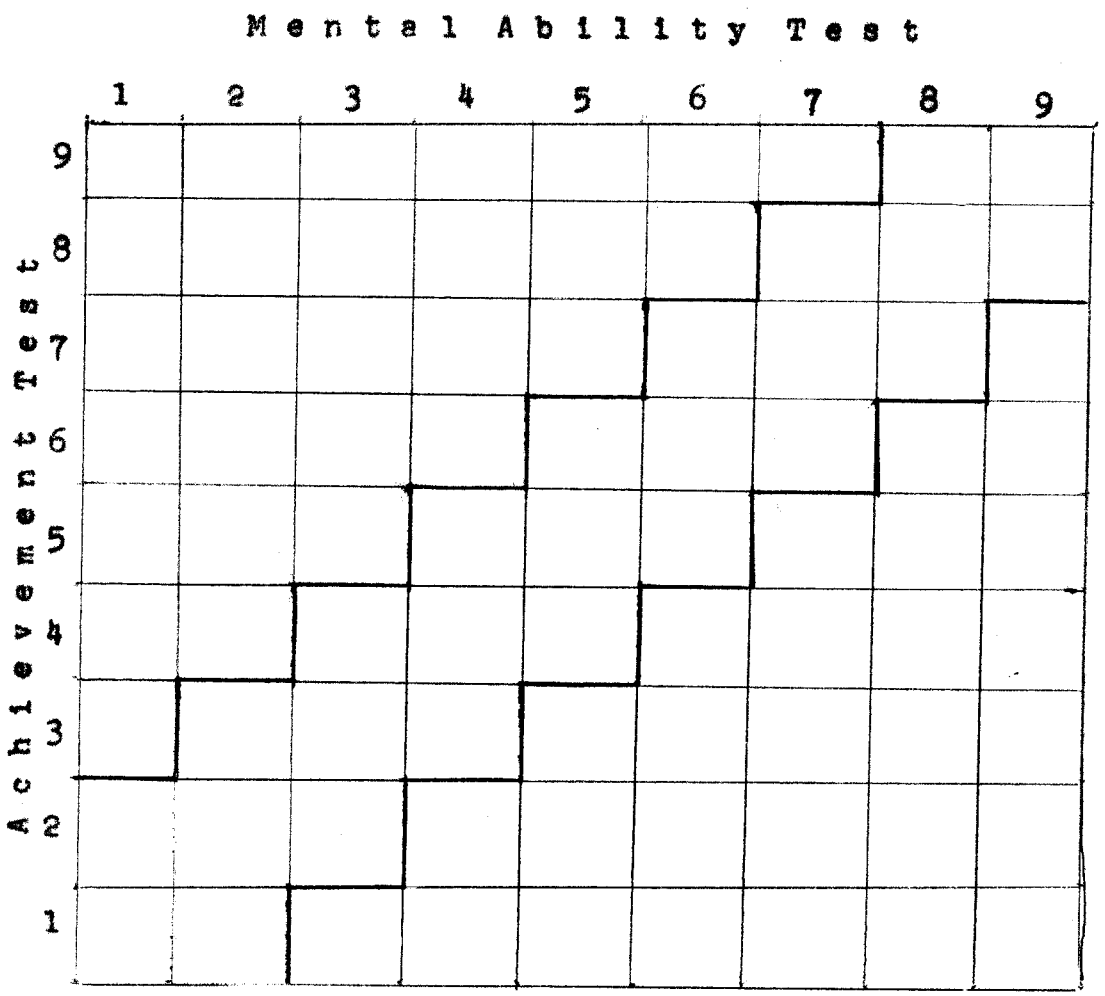
intelligence to measures of achievement only a difference of two or more stanines should be given consideration for further investigation, since smaller differences are likely to be measurement errors rather than true differences.

9. Scattergrams based on stanines can be developed. This can be done by preparing a grid with nine squares running horizontally and nine vertically or a total of eighty-one cells excluding the marginal cells used for identification and for summarizing the distributions. The horizontal axis should be labelled as mental ability from a stanine of one on the extreme left to a stanine of nine on the right. The vertical axis would be labelled as achievement with the bottom row being a stanine of one and the top row being a stanine of nine. Heavy zig-zag lines could then be drawn to lay off a band plus and minus stanine deviation from the lower-left to upper-right diagonal. After the data for pupils are tallied in the appropriate cell those cases falling outside the zig-zag lines

(above or below) should be investigated first.

TABLE XXIX

PROPOSED STANINE SCATTERGRAM



BIBLIOGRAPHY

I. BOOKS

Berry, Ruth and Beverly Wolf, An Epitaph For Vocational Guidence, New York, 1962.

DuBois, Philip A. and Edward Hackett, The Measurement and Evaluation of Over and Underachievement, Contract Monograph (816) St. Louis, 1962.

Germane, Charles, and Edith Germane, Personnel Work In High School, New York, 1941.

Goodenough, Florence, Mental Testing, New York, 1949.

Greene, Harry A., and Albert N. Jorgenson, The Use and Interpretation of Elementary School Tests, New York, 1937.

Guilford, J. P., Psychometric Methods, New York, 1954.

Harris, Chester W., editor, Encyclopedia of Educational Research, New York, 1960.

Kelley, Truman Lee, Interpretation of Educational Measurements, New York, 1927.

Lindquist, E. F., ed. Educational Measurement, Washington, D.C., 1951.

Loughin, Leonora, The Relative Achievement of Pupils of Two Public Schools, M.A. Thesis, University of Chicago, 1937.

McCall, W. A., How to Measure in Education, New York, 1922.

Noble, M. C., Practical Measurement for School Administrators, Scranton, 1939.

Pressey, Sidney and Luella Pressey, Introduction to the Use of Standardized Tests, New York, 1923.

Sherman, Charles E., Age and Grade Status Factors In Relation To Achievement In Children Grade 3 - 8, Unpublished Thesis Loyola University, Chicago, 1963.

Thorndike, Robert L., The Concepts of Over- and Underachievement, New York, 1963.

Thurstone, L. L., Multiple Factor Analysis, Chicago, 1947.

II. ARTICLES

Asher, William, "Statistical Problems of the Accomplishment Quotient," Journal of Experimental Education, XXX, March, 1962.

Beeson, M. F., and R. E. Tope, "E.Q. and A.Q. as Aids in Classification of Pupils," Journal of Educational Research, IX, April, 1924.

Bowman, Howard, "Techniques for Graphic Representation of Pupil Personnel Data to Indicate Individual Deviates and to Provide a Basis for More Adequate Guidance," Educational and Psychological Measurement, XII No. 3, 1952.

Chapman, J. Crosby, "The Unreliability of the Difference Between Intelligence and Educational Ratings," Journal of Educational Psychology, XIV, February, 1923.

Coleman, William and Edward Cureton, "Intelligence and Achievement The Jangle Fallacy," Educational and Psychological Measurement, Vol. 14, 1954.

Cory, Genevieve L., "A Study of Various Factors Which Influence the Use of the Accomplishment Quotient As A Measure of Teaching Efficiency," Journal of Educational Research, XXI, January, 1930.

Cureton, E. F., "The Accomplishment Quotient Technic," Journal of Experimental Education, March, 1937.

Durost, W. N. and G. A. Prescott, "An Improved Method of Comparing a Capacity Measure and An Achievement Meesure At the Elementary School Level," Educational Psychological Measurement, XII, 1952.

"Educational Measurement - Interpreting and Using the Results," The National Elementary Principal, XLI No. 2, November, 1961.

"Educational Measurement - Purposes and Techniques," The National Elementary Principal, XLI No. 1, September, 1961.

Engelhart, Max, "A Useful Technique for Analyzing Test Results," Chicago Schools Journal, May, 1962.

Madsen, I. N., "Interpreting Achievement in School in Terms of Intelligence," American School Board Journal, November, 1921.

McDonald, D. W., "Scatter Diagrams and Their Use," The School Guidance Worker, XVI, March, 1961.

McPhail, Andrew, "The Correlation Between I.Q. and A.Q.," School and Society, XVI, November, 1922.

Monroe, Walter S., "The Illinois Examination," Bureau of Educational Research Bulletin, No. 6 IX, October, 1921.

Morley, Clyde, "The Reliability of the Achievement Quotient," The Journal of Educational Psychology, XXI, May, 1930.

Murdock, A., "The A.Q., Finding and Using It," Teachers' College Record, XXIII, May, 1922.

Mygaard, P.H., "A Revised Accomplishment Quotient," Journal of Educational Research, XVIII, June, 1928.

Shanner, William, "New Concepts In Norms," Positive Values in the American Educational System, American Council on Education, Washington, D.C., 1959.

Stebbins, Rena and L. A. Pechstein, "Quotients I. E. and A," Journal of Educational Psychology, XIII, October, 1922.

Tiedeman, David and Charles McArthur, "Over and Underachievement If Any," 13th Yearbook National Council of Measurements Used In Education, New York, 1956.

Torgerson, T. L. and Irene Schuman, "The Variability of the Accomplishments of Pupils of the Same Mental Level," Journal of Educational Research, XI, February, 1925.

Wilson, William, "The Misleading Accomplishment Quotient," Journal of Educational Research, VII, January, 1928.

Wolfe, Dael, "Factor Analysis to 1940," Psychometric Monograph No. 3, 1946.

III. TEST HANDBOOKS

A Guidance Handbook for the Adjustment Services in the Elementary Schools, Chicago Public Schools, Chicago, 1954.

Expected Achievement Grade Placement Tables, Division of Instructional Services, Denver Public Schools, 1962.

Expected Achievement Grade Placement Tables, Dividion of Instructional Services, Los Angeles City School Districts, 1955.

Group Testing Bulletin II Testing Office, Department of Guidance and Counseling, Tucson Public Schools, 1962.

Psychologists' Handbook, Bureau of Child Study, 1954.

Test Handbook for the Elementary Schools, Research Department, Wichita Public Schools, 1962.

The County-Wide Group Testing Program, Dade County, Florida Public Schools, 1960.

IV. TESTS

Durost, Walter N., et al., Manual: Metropolitan Achievement Tests, Chicago and New York, 1959.

Sullivan, Elizabeth, et al., Manual: California Short-Form Test of Mental Maturity, Los Angeles, 1957.

APPENDIX I-A

LIST OF CITIES AND THEIR RESPONSE TO A QUESTIONNAIRE REGARDING CURRENT PRACTICES OF RELATING MENTAL ABILITY AND ACHIEVEMENT TEST DATA

Cities	Response: Yes	Response: No
Akron, Ohio	x	
Albuquerque, N. Mex.	x	
Atlanta, Ga.	x	
Baltimore, Md.	x	
Birmingham, Ala.	x	
Boston, Mass.	x	
Buffalo, N.Y.		x
Charlotte, N. Car.	x	
Cincinnati, Ohio	x	
Cleveland, Ohio	x	
Columbus, Ohio	x	
Dallas, Tex.		x
Dayton, Ohio	x	
Denver, Colo.	x	
Des Moines, Iowa	x	
Detroit, Mich.	x	
Fort Worth, Tex.		x
Houston, Tex.	x	
Indianapolis, Ind.	x	
Jacksonville, Fla.	x	
Jersey City, N.J.	x	
Kansas City, Mo.	x	
Long Beach, Calif.		x
Los Angeles, Calif.	x	
Louisville, Ky.		x

Cities	Response: Yes	Response: No
Memphis, Tenn.	x	
Miami, Fla.	x	
Milwaukee, Wis.	x	
Minneapolis, Minn.	x	
Newark, N.J.	x	
New Orleans, La.	x	
New York, N.Y.	x	
Norfolk, Va.	x	
Oakland, Calif.	x	
Oklahoma City, Okla.	x	
Omaha, Nebr.	x	
Philadelphia, Pa.	x	
Phoenix, Ariz.	x	
Pittsburgh, Pa.	x	
Portland, Ore.	x	
Providence, R.I.		x
Richmond, Va.	x	
Rochester, N.Y.	x	
St. Louis, Mo.	x	
St. Paul, Minn.	x	
San Antonio, Tex.	x	
San Diego, Calif.	x	
San Francisco, Calif.	x	
Seattle, Wash.		x
Syracuse, N.Y.	x	
Tampa, Fla.	x	
Toledo, Ohio	x	
Tucson, Ariz.	x	
Tulsa, Okla.	x	
Washington, D.C.	x	
Wichita, Kan.	x	

Total - 56 cities

Appendix I-B

December 12, 1962

Dear Sir:

As a graduate student at Loyola University of Chicago, I am interested in investigating the various methods and practices which are utilized in interpreting measures of school achievement in relation to measures of mental maturity or intelligence. In order to obtain data about current practices, the enclosed questionnaire has been prepared and is being sent to a number of large school systems.

This study is being done under the supervision of Dr. Max Engelhart, Director of Student Examinations, Chicago City Junior College and a member of the graduate school faculty of Loyola. The information obtained as a result of this questionnaire will in no way be used to make judgments about your school system. If this request meets with your approval, please direct the questionnaire to the person in your system best qualified to answer the questions with reference to your testing program.

Thank you for your cooperation.

Yours truly,

/jr
enc. 3

John A. Russell
2812 West 100th Street
Evergreen Park 42, Illinois

QUESTIONNAIRE

PRACTICES OF INTERPRETING MEASURES OF ACHIEVEMENT IN RELATION TO MEASURES OF INTELLIGENCE

1. Do you have a city-wide standardized testing program? (Yes___)
(No___)

If yes, please give the following information:

Grade Level	Mental Ability Test or Tests (Name)	Achievement Test or Tests (Name & Kind)

2. What methods are utilized to relate mental ability test results to achievement test results for individual students?

3. What provisions are made to have uniform analysis?

_____

_____

_____

_____

_____

_____

_____

4. Have printed or duplicated materials describing the techniques and methods to be followed in relating mental ability test results to achievement test results been developed?

Yes_____ No_____

(If yes, please send a copy of these materials when you return this questionnaire.)

Name of person preparing the questionnaire _____

Title _____

Your co-operation in preparing this questionnaire is sincerely appreciated. PLEASE RETURN it in the attached self-addressed envelope to:

John A. Russell
2812 West 100th Street
Evergreen Park 42, Illinois

APPENDIX II-A

WORK SHEET FOR SIXTH GRADE FACTOR ANALYSIS

Test	1	2	3	4	5	6	7	8	9	10	11	Check Sum
1	(.84)	.83	.75	.74	.75	.64	.70	.75	.68	.84	.72	8.24
2	.83	(.82)	.66	.74	.77	.65	.72	.76	.71	.82	.71	8.19
3	.75	.66	(.67)	.67	.65	.60	.63	.60	.52	.63	.60	6.98
4	.74	.74	.67	(.76)	.76	.70	.73	.67	.67	.70	.70	7.84
5	.75	.77	.65	.76	(.75)	.66	.71	.72	.70	.75	.68	7.90
6	.64	.65	.60	.70	.66	(.83)	.83	.58	.61	.60	.66	7.36
7	.70	.72	.63	.73	.71	.83	(.74)	.66	.70	.69	.74	7.85
8	.75	.76	.60	.67	.72	.58	.66	(.80)	.73	.80	.65	7.72
9	.68	.71	.52	.67	.70	.61	.70	.73	(.75)	.75	.69	7.51
10	.84	.82	.63	.70	.75	.60	.69	.80	.75	(.70)	.70	7.98
11	.72	.71	.60	.70	.68	.66	.74	.65	.69	.70	(.74)	7.59
Sum of Rows	8.24	8.19	6.98	7.84	7.90	7.36	7.85	7.72	7.51	7.98	7.59	85.16
	.893	.887	.756	850	.856	.798	.851	.837	.814	.865	.822	9.228

Code for Tests

Metropolitan Battery

1. Word Knowledge
2. Reading
3. Spelling
4. Language
5. Language Study Skills
6. Arithmetic Computation
7. Arithmetic Problem Solving
8. Social Studies Information
9. Social Studies Study Skills
10. Science

California Test Mental Maturity

11. Total

WORK SHEET FOR EIGHTH GRADE FACTOR ANALYSIS

Test	1	2	3	4	5	6	7	8	9	10	11	Check Sum
1	(.85)	.85	.65	.73	.73	.65	.70	.75	.69	.76	.71	8.07
2	.85	(.78)	.64	.75	.74	.68	.72	.78	.72	.75	.69	8.10
3	.65	.64	(.69)	.69	.61	.58	.55	.55	.48	.46	.50	6.40
4	.73	.75	.69	(.74)	.74	.74	.73	.66	.70	.64	.66	7.78
5	.73	.74	.61	.74	(.69)	.69	.69	.63	.68	.64	.61	7.45
6	.65	.68	.58	.74	.69	(.84)	.84	.64	.70	.62	.67	7.65
7	.70	.72	.55	.73	.69	.84	(.77)	.67	.77	.71	.73	7.88
8	.75	.78	.55	.66	.63	.64	.67	(.74)	.67	.74	.63	7.46
9	.69	.72	.48	.70	.68	.70	.77	.67	(.70)	.70	.70	7.51
10	.76	.75	.46	.64	.64	.62	.71	.74	.70	(.67)	.67	7.36
11	.71	.69	.50	.66	.61	.67	.73	.63	.70	.67	(.73)	7.30
Sum of Rows	8.07	8.10	6.40	7.78	7.45	7.65	7.88	7.46	7.51	7.36	7.30	82.96
	.886	.889	.703	.854	.818	.840	.865	.819	.825	.809	.801	9.108

Code for Tests

Metropolitan Battery

1. Word Knowledge
2. Reading
3. Spelling
4. Language
5. Language Study Skills
6. Arithmetic Computation
7. Arithmetic Problem Solving
8. Social Studies Information
9. Social Studies Study Skills
10. Science

California Test Mental Maturity

11. Total

APPROVAL SHEET

The dissertation submitted by John A. Russell has been read and approved by five members of the Department of Education.

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated, and that the dissertation is now given final approval with reference to content, form, and mechanical accuracy.

The dissertation is therefore accepted in partial fulfillment of the requirements for the degree of Doctor of Education.

June 5, 1964
Date

Max D. Engelhart
Signature of Adviser